

**Titre:** Application de méthodes de programmation dynamique  
stochastique au problème de planification de la production  
d'hydroélectricité  
**Title:**

**Auteur:** Mouad Faik  
**Author:**

**Date:** 2015

**Type:** Mémoire ou thèse / Dissertation or Thesis

**Référence:** Faik, M. (2015). Application de méthodes de programmation dynamique  
stochastique au problème de planification de la production d'hydroélectricité  
**Citation:** [Mémoire de maîtrise, École Polytechnique de Montréal]. PolyPublie.  
<https://publications.polymtl.ca/1832/>

 **Document en libre accès dans PolyPublie**  
Open Access document in PolyPublie

**URL de PolyPublie:**  
PolyPublie URL: <https://publications.polymtl.ca/1832/>

**Directeurs de  
recherche:** Michel Gendreau, & Stéphane Krau  
**Advisors:**

**Programme:** Maîtrise en mathématiques appliquées  
**Program:**

UNIVERSITÉ DE MONTRÉAL

APPLICATION DE MÉTHODES DE PROGRAMMATION DYNAMIQUE  
STOCHASTIQUE AU PROBLÈME DE PLANIFICATION DE LA PRODUCTION  
D'HYDROÉLECTRICITÉ

MOUAD FAIK  
DÉPARTEMENT DE MATHÉMATIQUES ET DE GÉNIE INDUSTRIEL  
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION  
DU DIPLOME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES  
(MATHÉMATIQUES APPLIQUÉES)  
AOÛT 2015

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

APPLICATION DE MÉTHODES DE PROGRAMMATION DYNAMIQUE  
STOCHASTIQUE AU PROBLÈME DE PLANIFICATION DE LA PRODUCTION  
D'HYDROÉLECTRICITÉ

présenté par : FAIK Mouad

en vue de l'obtention du diplôme de : Maîtrise ès sciences appliquées

a été dûment accepté par le jury d'examen constitué de :

M. AUDET Charles, Ph. D., président

M. GENDREAU Michel, Ph. D., membre et directeur de recherche

M. KRAU Stéphane, Ph. D., membre et codirecteur de recherche

M. BASTIN Fabian, Doctorat, membre

## DÉDICACE

*À mes frères Mohammed et Hicham  
qui m'encourageaient toujours  
à persévérer dans mes études  
et viser plus haut.  
À toute ma famille...*

## REMERCIEMENTS

Je tiens d’abord à remercier mon directeur de recherche Michel Gendreau pour son support tout au long de ma maîtrise. Je remercie en son nom le fond de recherche CRSNG-Hydro Québec pour avoir financé ce projet. Mes remerciements vont également à mon codirecteur de recherche Stéphane Krau pour son implication et sa disponibilité. J’exprime aussi mes remerciement à Gréegory Émiel, consultant en modélisation à Hydro-Québec, pour ses suggestions fructueuses et les échanges scientifiques que nous avons eus. Je remercie les membres du jury qui ont enrichi ce mémoire par leurs commentaires pertinents.

Je souhaite exprimer ma gratitude à l’Institut national polytechnique de Grenoble, plus particulièrement l’École nationale supérieure d’informatique et de mathématiques appliquées de Grenoble (ENSIMAG) pour m’avoir donné l’opportunité de faire un double diplôme à l’École Polytechnique de Montréal. Je remercie tous les professeurs des deux écoles pour la richesse de leurs enseignements.

Je remercie tout le personnel du CIRRELT, en particulier Serge Bisailon, pour l’aide quotidienne qu’il apporte aux étudiants.

Je remercie tous mes amis sans exception pour les moments inoubliables que j’ai partagés avec eux que ce soit en étudiant ou en jouant ou en voyageant...

Je tiens à exprimer ma vive reconnaissance envers mes parents, mes frères et sœurs ainsi que mes neveux et nièces pour leur support moral continu et leur encouragement durant tout mon parcours académique. Un grand merci à ma fiancée Ola pour son aide précieuse dans l’édition de plusieurs graphiques utilisés dans ce mémoire ainsi que pour ses lectures et corrections de mon texte.

## RÉSUMÉ

La planification de la production d'hydroélectricité est un problème complexe qui se situe au croisement de plusieurs branches des mathématiques appliquées. La prise de décisions optimales ou quasi-optimales passe par l'optimisation, parfois déterministe et en nombre entiers (planification court-terme), parfois stochastique et continue (planification moyen- et long-terme). La maîtrise de l'aspect stochastique des apports hydriques naturels et/ou des prix de l'énergie passe par l'utilisation et le développement de modèles statistiques probabilistes. La modélisation des fonctions de production des centrales hydroélectriques passe par l'optimisation non-linéaire et des méthodes de calcul numériques.

Dans ce mémoire, deux volets sont étudiés : méthodes d'approximation concave des fonctions de production des centrales hydroélectriques et méthodes d'optimisation stochastique appliquées au problème de planification moyen-terme de la production d'hydroélectricité sous incertitude portant sur les apports hydriques naturels.

Les fonctions de production des centrales hydroélectriques dépendent du débit turbiné et de la hauteur de chute. Ces fonctions ne sont ni linéaires ni même concaves. Pour les incorporer à un modèle mathématique linéaire de maximisation, il faut au préalable les approximer par des fonctions concaves linéarisables. Pour ce faire, on propose deux méthodes : *approximation de la fonction objectif par enveloppe concave* et *approximation par une fonction concave linéaire par morceaux*. La dernière étant en deux variantes : construction d'hyperplans d'ajustement avec le critère des moindres carrés ou le critère de minimisation de l'erreur maximale par rapport à la vraie fonction de production. On conclut que les deux méthodes sont à peu près équivalentes en matière d'erreur-type mais que la deuxième méthode est meilleure sur le temps d'exécution des modèles de planification de la production utilisés.

Le problème de planification moyen-terme de la production d'hydroélectricité sous incertitude peut être approché et résolu par différentes méthodes. Les méthodes de programmation dynamique stochastique restent les plus rigoureuses en matière de formulation et d'algorithmes de résolution et de prise en compte de l'incertitude. On se propose d'appliquer deux méthodes basées sur la programmation dynamique stochastique : la programmation dynamique duale stochastique (SDDP) et une méthode de programmation dynamique stochastique approchée (ASDP) dont l'idée centrale est de discrétiser l'espace des états en hyper-rectangles. On compare les deux méthodes sur un cas d'étude réel : La Grande Rivière, la rivière la plus importante du parc hydroélectrique du Québec.

## ABSTRACT

The hydro-power generation planning is a complex problem implicating different fields of applied mathematics. Optimal or almost-optimal decisions are taken after optimization. This optimization is sometimes deterministic with integer variables (short-term planning), sometimes it is stochastic and continuous (mid or long-term planning). In order to master the randomness of parameters such as inflows or spot prices one must use and develop statistical and probabilistic models. Hydro-plants' power generation functions are modeled using non-linear optimization, numerical calculating methods. etc.

This thesis focuses on two main subjects: concavation methods of hydro-plants' power generation functions and stochastic optimization methods applied to a mid-term hydro-power generation planning problem under uncertainty that lies on natural inflows.

Hydro-plants' power generation functions depend on the turbined outflow and water head, they are consequently not linear nor concave. These functions must be approximated by concave, piecewise linear functions, so they can fit into a linear mathematical programming model. There are two methods to do so: *approximation of the objective function by a concave hull* and *approximation by a concave piecewise linear function*. The second approach has two variants: fitting hyperplanes construction with the least squares criterion or the criterion of minimizing the maximal error to the real generation function. We conclude that these two methods are almost equivalent according to their root-mean square error, but the second method has less impact on planning algorithms' running time.

There are many methods that can be used to solve the mid-term hydro-power generation planning under uncertainty. Stochastic dynamic programming-based methods are still the most rigorous with respect to their formulation, the solution algorithms used and the way in which uncertainty is represented. Two of these methods are compared in this thesis: Stochastic Dual Dynamic Programming (SDDP) and a variant of Approximate Stochastic Dynamic Programming (ASDP), which is mainly based on a state space discretization into hyper-rectangles. These two methods are compared on a real case-study: La Grande Rivière, the most important river in Quebec hydroelectricity system.

## TABLE DES MATIÈRES

DÉDICACE . . . . .	iii
REMERCIEMENTS . . . . .	iv
RÉSUMÉ . . . . .	v
ABSTRACT . . . . .	vi
TABLE DES MATIÈRES . . . . .	vii
LISTE DES TABLEAUX . . . . .	x
LISTE DES FIGURES . . . . .	xi
LISTE DES SIGLES ET ABRÉVIATIONS . . . . .	xiii
LISTE DES ANNEXES . . . . .	xiv
CHAPITRE 1 INTRODUCTION . . . . .	1
1.1 Généralités . . . . .	1
1.2 Contexte : le Québec, un immense potentiel hydroélectrique . . . . .	3
1.3 Objectifs . . . . .	5
1.4 Contenu du mémoire . . . . .	5
CHAPITRE 2 REVUE DE LITTÉRATURE . . . . .	7
2.1 Gestion de réservoirs sous incertitude . . . . .	7
2.2 Résolution par réduction à un arbre de scénarios . . . . .	11
2.3 Résolution par programmation dynamique stochastique approchée . . . . .	12
2.4 Résolution par programmation dynamique duale stochastique . . . . .	13
2.5 Fonctions de production des centrales hydroélectriques . . . . .	15
CHAPITRE 3 ÉLÉMENTS DE MODÉLISATION . . . . .	17
3.1 Infrastructures et installations hydroélectriques . . . . .	17
3.1.1 Réservoir . . . . .	18
3.1.2 Centrale hydroélectrique . . . . .	19
3.1.3 Ouvrage d'évacuation . . . . .	20



3.2	Fonction de production d'une centrale hydroélectrique . . . . .	20
3.2.1	Cas d'un groupe turbine-alternateur . . . . .	21
3.2.2	Cas d'une centrale . . . . .	23
3.2.3	Modélisation pour la planification moyen-terme de la production . . .	24
3.2.4	Approximation concave des fonctions de production . . . . .	27
3.3	Apports hydriques naturels . . . . .	31
3.3.1	Modèles auto-régressifs périodiques . . . . .	31
3.3.2	Estimation des paramètres du modèle PAR(1) . . . . .	33
3.3.3	Génération d'un processus stochastique PAR(1) discrétisé . . . . .	33
3.4	Classes de demande . . . . .	36
CHAPITRE 4 PROBLÈME DE PLANIFICATION DE LA PRODUCTION D'HY-		
DROÉLECTRICITÉ . . . . .		37
4.1	Notation . . . . .	37
4.2	Formulation du problème . . . . .	40
4.3	Description du problème . . . . .	41
4.3.1	Objectifs . . . . .	41
4.3.2	Contraintes . . . . .	42
CHAPITRE 5 MÉTHODES DE RÉOLUTION PAR PROGRAMMATION DYNA-		
MIQUE STOCHASTIQUE . . . . .		48
5.1	Notation . . . . .	48
5.2	Reformulation du problème : schéma de programmation dynamique . . . .	49
5.3	Programmation dynamique stochastique et politique de gestion . . . . .	50
5.4	Concavité des fonctions de bénéfice futur . . . . .	51
5.4.1	Concavité par rapport à l'état <i>volume</i> . . . . .	51
5.4.2	Concavité par rapport à l'état <i>apport passé</i> . . . . .	52
5.5	Représentation des fonctions de bénéfice futur . . . . .	53
5.6	Méthode de programmation dynamique stochastique exacte . . . . .	57
5.7	Méthode de programmation dynamique duale stochastique (SDDP) . . . .	59
5.7.1	Passe <i>vers l'avant</i> . . . . .	60
5.7.2	Passe <i>vers l'arrière</i> . . . . .	60
5.7.3	L'algorithme SDDP . . . . .	61
5.7.4	Critère d'arrêt . . . . .	62
5.7.5	Convergence . . . . .	64
5.8	Méthode de programmation dynamique stochastique approchée (ASDP) . . .	64
5.8.1	Discrétisation adaptée de l'espace d'état <i>volume</i> . . . . .	65

5.8.2	Discrétisation adaptée simplifiée de l'espace d'état <i>volume</i> . . . . .	66
5.8.3	L'algorithme ASDP . . . . .	67
5.8.4	Critère d'arrêt . . . . .	69
5.8.5	Convergence . . . . .	69
CHAPITRE 6	ANALYSE D'UN CAS D'ÉTUDE : LA GRANDE RIVIÈRE . . . . .	70
6.1	Description du cas d'étude . . . . .	70
6.1.1	Caractéristiques des installations hydroélectriques . . . . .	70
6.1.2	Demande en électricité . . . . .	73
6.2	Comparaison de différentes modélisations des fonctions de production des centrales . . . . .	73
6.2.1	Modélisation par enveloppe concave . . . . .	74
6.2.2	Modélisation par une fonction concave affine par morceaux . . . . .	74
6.2.3	Conclusion . . . . .	77
6.3	Planification moyen-terme de la production d'hydroélectricité : SDDP vs ASDP . . . . .	78
6.3.1	Évolution des bornes . . . . .	78
6.3.2	Raffinement de la politique . . . . .	80
6.3.3	Échantillonnage de l'espace des états . . . . .	80
6.3.4	Gestion des réservoirs et de la production de LGR . . . . .	84
6.4	Retour sur les fonctions de production . . . . .	89
CHAPITRE 7	CONCLUSION . . . . .	90
7.1	Synthèse des travaux . . . . .	90
7.2	Limitations de la solution proposée . . . . .	91
7.3	Améliorations futures . . . . .	91
RÉFÉRENCES	. . . . .	92
ANNEXES	. . . . .	98

## LISTE DES TABLEAUX

Tableau 1.1	Pourcentage de l'hydroélectricité dans la production totale . . . . .	1
Tableau 5.3	Complexité du schéma SDP exact . . . . .	59
Tableau 6.1	Caractéristiques des réservoirs de LGR . . . . .	70
Tableau 6.2	Caractéristiques des centrales de LGR . . . . .	71
Tableau 6.3	Caractéristiques des ouvrages d'évacuation de LGR . . . . .	71
Tableau 6.4	Taille des grilles utilisées . . . . .	74
Tableau 6.5	Erreur-type de la production estimée (cas de <b>EnvConc</b> ) . . . . .	74
Tableau 6.6	Ordre de grandeur du temps d'exécution de l'algorithme de partitionnement . . . . .	77
Tableau 6.7	Volumes initiaux des réservoirs à capacité positive . . . . .	78
Tableau 6.8	Nombre d'hyperplans de l'approximation de la fonction de production	89

## LISTE DES FIGURES

Figure 1.1	Source d'électricité au Québec . . . . .	2
Figure 1.2	Hydroélectricité du Québec . . . . .	4
Figure 2.1	Processus décisionnel à deux étapes . . . . .	8
Figure 2.2	Procédure d'optimisation stochastique implicite (Labadie (2004)) . . . . .	9
Figure 2.3	Procédure d'optimisation stochastique explicite (Labadie (2004)) . . . . .	10
Figure 2.4	Un arbre de scénarios . . . . .	11
Figure 3.1	Processus de production de l'hydroélectricité . . . . .	17
Figure 3.2	Barrage de la Trenché (Québec) . . . . .	18
Figure 3.3	Schéma de base d'un réservoir . . . . .	18
Figure 3.4	Centrale Robert-Bourassa, une centrale souterraine à réservoir (Québec) . . . . .	19
Figure 3.5	Centrale de Carillon, une centrale au fil de l'eau (Québec) . . . . .	19
Figure 3.6	Évacuateur de crues du barrage Robert-Bourassa (Québec) . . . . .	20
Figure 3.7	Rendement d'un groupe turbine-alternateur . . . . .	21
Figure 3.8	Exemple de fonction débit-puissance . . . . .	22
Figure 3.9	Transformation par similitude de la fonction débit-puissance . . . . .	23
Figure 3.10	Exemple de configuration physique d'une centrale . . . . .	23
Figure 3.11	Exemple de fonction de production à hauteur de référence (Centrale à deux groupes) . . . . .	25
Figure 3.12	Exemple de fonction de production globale . . . . .	26
Figure 3.13	Historique des apports hydriques (53 années) . . . . .	32
Figure 3.14	Génération naïve du processus stochastique discrétisé . . . . .	34
Figure 3.15	Génération efficace du processus stochastique discrétisé . . . . .	35
Figure 3.16	Classes de demande . . . . .	36
Figure 5.1	Deux scénarios partageant les mêmes apports jusqu'à la semaine $t$ . . . . .	50
Figure 5.2	Illustration d'une fonction de bénéfice futur et ses hyperplans supports . . . . .	55
Figure 5.3	Approximation d'une fonction de bénéfice futur par des hyperplans coupants . . . . .	56
Figure 5.4	Calcul analytique de l'erreur sur une arête d'un hyper-rectangle . . . . .	67
Figure 6.1	Schéma de La Grande Rivière . . . . .	72
Figure 6.2	Demande électrique . . . . .	73
Figure 6.3	Relation entre, (1) la discrétisation de l'espace débit-hauteur de chute, (2) l'erreur-type ( $RMSE_G$ ), et (3) le nombre d'hyperplans approximant la fonction de production (cas de la centrale Caniapiscau) . . . . .	75

Figure 6.4	Relation entre, (1) la discrétisation de l'espace débit-hauteur de chute, (2) l'erreur-type ( $RMSE_G$ ), et (3) le nombre d'hyperplans approximant la fonction de production (cas de la centrale La Grande <sub>4</sub> ) . . . . .	75
Figure 6.5	Relation entre, (1) la discrétisation de l'espace débit-hauteur de chute, (2) l'erreur-type ( $RMSE_G$ ), et (3) le nombre d'hyperplans approximant la fonction de production (cas de la centrale La Grande <sub>3</sub> ) . . . . .	76
Figure 6.6	Relation entre, (1) la discrétisation de l'espace débit-hauteur de chute, (2) l'erreur-type ( $RMSE_G$ ), et (3) le nombre d'hyperplans approximant la fonction de production (cas de la centrale La Grande <sub>2</sub> ) . . . . .	76
Figure 6.7	(a) évolution des bornes inférieure et supérieure ; (b) zoom sur la convergence. . . . .	79
Figure 6.8	Amélioration de la politique sur 1000 scénarios fixes . . . . .	80
Figure 6.9	Distribution des coupes de Benders sur les différents cas d'apport . . . . .	81
Figure 6.10	Transitions dans le processus d'apports PAR(1) discrétisé . . . . .	82
Figure 6.11	Volumes échantillonnés par les méthodes SDDP et ASDP . . . . .	83
Figure 6.12	Scénarios de simulation post-politique . . . . .	84
Figure 6.13	Diagramme en boîte du volume moyen simulé du réservoir Caniapiscau . . . . .	85
Figure 6.14	Diagramme en boîte du volume moyen simulé du réservoir La Grande <sub>4</sub> . . . . .	85
Figure 6.15	Diagramme en boîte du volume moyen simulé du réservoir La Grande <sub>3</sub> . . . . .	86
Figure 6.16	Diagramme en boîte du volume moyen simulé du réservoir La Grande <sub>2</sub> . . . . .	86
Figure 6.17	Diagramme en boîte du débit déversé simulé de l'ouvrage d'évacuation Caniapiscau . . . . .	87
Figure 6.18	Diagramme en boîte de la puissance produite simulée durant la classe pointe . . . . .	88
Figure 6.19	Diagramme en boîte de la puissance produite simulée durant la classe moyenne . . . . .	88
Figure A.1	Une fonction concave linéaire par morceaux . . . . .	98

## LISTE DES SIGLES ET ABRÉVIATIONS

### Sigles et abréviations

SDP	Stochastic Dynamic Programming
SDDP	Stochastic Dual Dynamic Programming
ASDP	Approximate Stochastic Dynamic Programming
PAR	Periodic Autoregressive
RMSE	Root Mean Square Error
LGR	La Grande Rivière

## LISTE DES ANNEXES

ANNEXE A	REPRÉSENTATION DES FONCTIONS CONCAVES LINÉAIRES PAR MORCEAUX DANS LES MODÈLES LINÉAIRES . . . . .	98
----------	--	----

## CHAPITRE 1 INTRODUCTION

### 1.1 Généralités

L'hydroélectricité est une source d'énergie renouvelable qui constitue une grande part de l'électricité générée dans le monde. Le tableau 1.1<sup>1</sup> donne quelques statistiques sur l'utilisation de cette source :

Tableau 1.1 Pourcentage de l'hydroélectricité dans la production totale

Pays	% de l'hydroélectricité
Norvège	96.7
Brésil	75.2
Venezuela	64.8
Canada	60.0
Suède	47.5
Chine	17.5
Russie	15.6
Inde	11.2
Japon	8.1
États unis	7.0
Reste des pays	14.0
Tout le monde	16.5

Plusieurs aspects de l'hydroélectricité sont particulièrement attrayants :

- il s'agit d'une des sources d'énergie les plus propres : « une centrale hydroélectrique émet 10 grammes de dioxyde de carbone (CO<sub>2</sub>) par kilowattheure produit, autant que son équivalent éolien et quatre fois moins que son équivalent solaire avec panneaux solaires photovoltaïques. Une centrale hydroélectrique produit 100 fois moins de GES<sup>2</sup> qu'une centrale au charbon. »<sup>3</sup> ;
- peu de coûts de production, l'eau étant gratuitement exploitée, au contraire de l'électricité produite du nucléaire ou du gaz naturel etc ;

1. Source : <http://www.iea.org> Agence internationale de l'énergie (Statistiques de 2012)

2. Gaz à effet de serre.

3. Source : <http://www.hydroquebec.ca>



- présente peu de danger dans son exploitation, contrairement à l'énergie nucléaire (le séisme et tsunami de 2011 au Japon dont les conséquences ne sont toujours pas complètement contrôlées est encore très récent).

Au Québec, comme le montre la figure 1.1<sup>4</sup>, l'hydroélectricité est la source par excellence d'énergie :

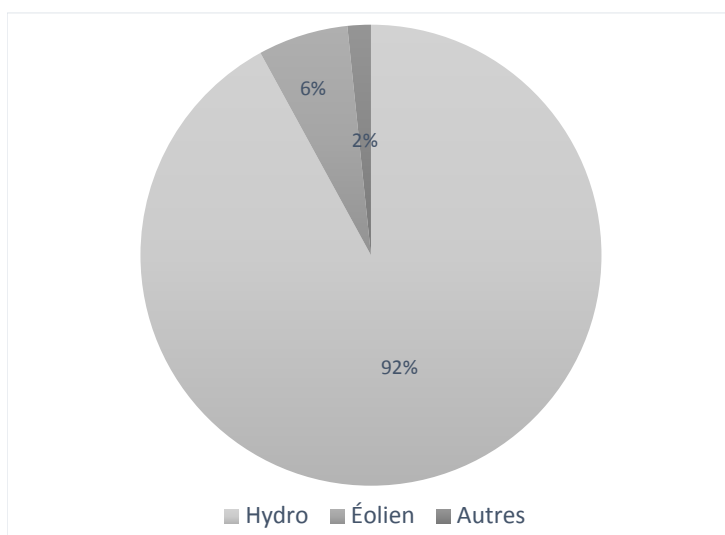


Figure 1.1 Source d'électricité au Québec

Le tableau et la figure précédents montrent le poids de l'hydroélectricité dans le monde, en particulier dans les systèmes à dominance hydroélectrique tels que le Norvège, le Brésil et le Québec. D'où l'importance de développer des modèles et méthodes de programmation mathématique pour représenter et résoudre le problème de gestion des ressources hydriques.

Les objectifs de la planification de la production d'électricité dans un système à dominance hydroélectrique peuvent varier selon son pourcentage d'énergie hydroélectrique, la nature du marché dont il fait partie, etc. On trouve les objectifs phares suivants :

- **Assurer la durabilité des ressources hydriques** : cet objectif est partagé par tous les différents systèmes ;
- **Minimiser les coûts de recours aux autres sources d'énergie** : cet objectif est fortement présent dans les systèmes où il y a un recours important aux autres sources d'énergie (thermique, nucléaire, etc) plus coûteuses, comme par exemple au Brésil ;

4. Adapté du rapport annuel d'Hydro-Québec de 2014.

- **Maximiser les bénéfices** : cet objectif est présent dans les systèmes dérégulés : système où l'échange (Import/Export) d'électricité est libéralisé comme par exemple en Norvège.

En pratique, le problème de la planification de la production hydroélectrique est approché par un processus décisionnel à plusieurs niveaux d'horizon de temps. On distingue habituellement les niveaux suivants :

- **Planification long-terme** : elle se fait sur un horizon de 3 ans et plus, fournit des objectifs finaux pour la planification moyen-terme (par exemple, les valeurs marginales de l'eau dans chaque réservoir à la fin de l'horizon) et peut aussi donner des indications sur les investissements futurs en infrastructure. Elle peut incorporer un modèle de changements climatiques ;
- **Planification moyen-terme** : elle se fait sur un horizon de 1 à 2 ans avec un pas de temps hebdomadaire, fournit des estimations des valeurs de l'eau à chaque fin de pas d'horizon et des entrées pour la planification court-terme ;
- **Planification court-terme** : elle se fait sur un horizon de quelques jours avec un pas horaire, fournit un plan de production détaillé qui sert à guider les opérations.

Ce processus décisionnel en trois niveaux fournit, si l'on peut ainsi dire, des « macro »-décisions qui vont piloter l'opération en temps réel qui vient ajouter d'autres variables de décisions, que l'on peut qualifier de « micro »-décisions, telles que la configuration à adopter pour chaque centrale (le nombre de groupes à engager et lesquels), la manière réelle dont l'électricité générée devra être transférée d'un nœud vers un autre du système de transport d'électricité.

## 1.2 Contexte : le Québec, un immense potentiel hydroélectrique

Le Québec est une terre abondante en ressources hydriques (500 000 lacs et 4500 rivières qui couvrent 12% de son territoire). Ses ressources sont exploitées par une infrastructure d'envergure : 656 barrages, 97 ouvrages régulateurs et 26 grands réservoirs d'une capacité de stockage de 175 TWh alimentant en eau 60 centrales hydroélectriques (d'une puissance installée pouvant aller jusqu'à 5 616 MW). La puissance totale installée pour ces centrales est de 35 364 MW. Ceci classe le système d'électricité du Québec parmi les plus écologiques au monde avec environ 99% d'hydroélectricité. La figure 1.2 montre les zones de concentration de la production, notamment la Grande Rivière (encadrée), et la manière avec laquelle elle est transportée via des liaisons électriques haute-tension.<sup>5</sup>

Cet important potentiel est géré par Hydro-Québec, une grande entreprise appartenant au

---

5. Source : <http://www.hydroquebec.ca>

## GRANDS ÉQUIPEMENTS

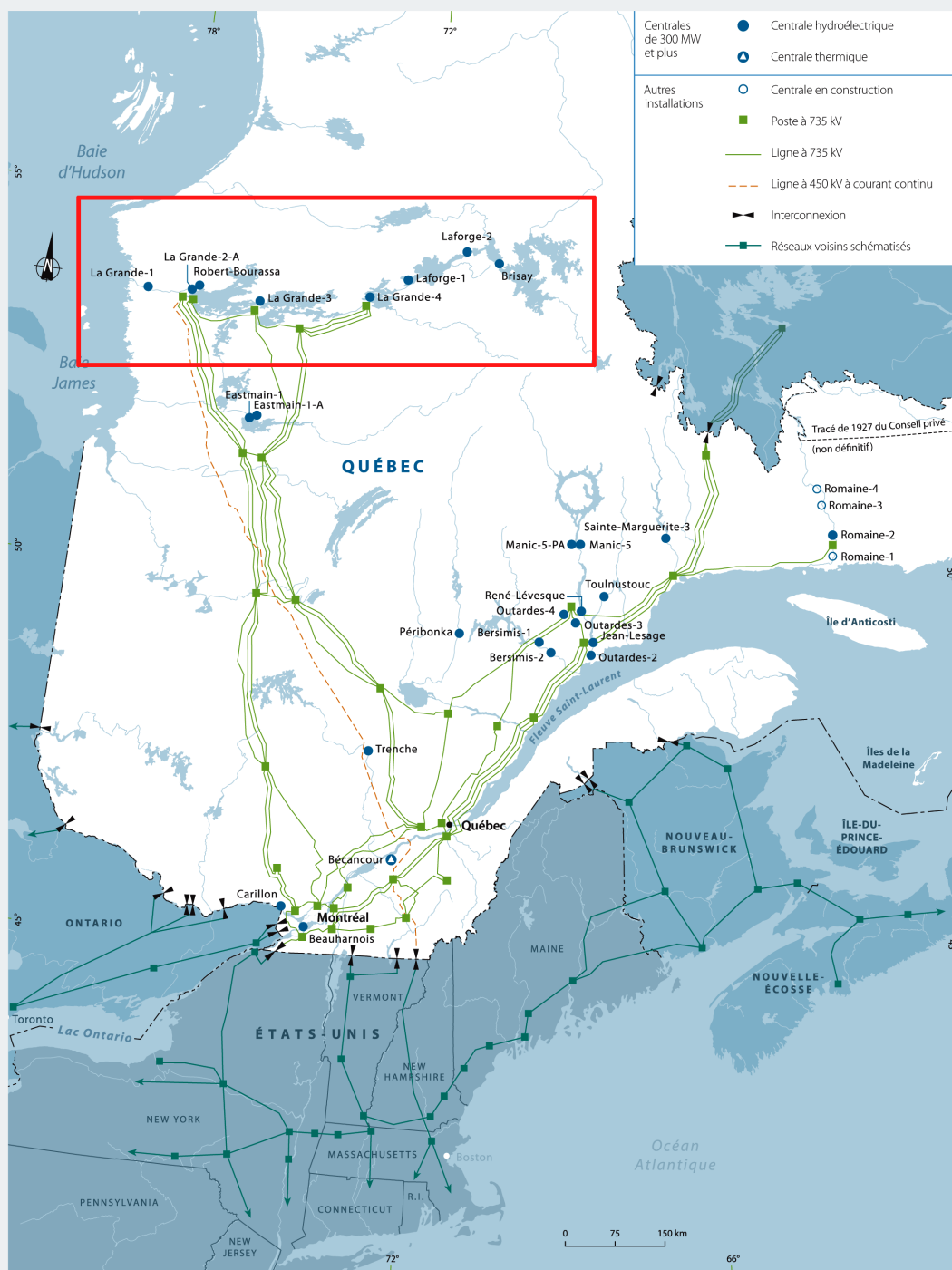


Figure 1.2 Hydroélectricité du Québec

gouvernement du Québec et leader mondial dans le domaine de l'hydroélectricité. Cette entreprise s'occupe de la gestion du parc hydroélectrique ainsi que des autres sources d'énergies exploitées (thermique, éolienne, etc.) en matière de production, de transport, d'investissement (notamment l'agrandissement du parc hydroélectrique) et de recherche et développement.

Parmi les enjeux majeurs de la gestion de la production de l'électricité au Québec on peut citer la maîtrise de l'aspect stochastique de certains phénomènes comme les apports naturels aux réservoirs et la demande. En d'autres termes, il y a besoin d'élaborer des modèles et méthodes de gestion capables de tenir compte, au mieux, de cette stochasticité.

### 1.3 Objectifs

L'objectif du présent mémoire est l'application de deux méthodes de programmation dynamique stochastique, à savoir la célèbre « Stochastic Dual Dynamic Programming » (SDDP), et une nouvelle méthode de programmation dynamique stochastique approchée (ASDP), à la planification moyen-terme de la production d'hydroélectricité. Pour atteindre cet objectif final, les étapes suivantes sont déterminantes :

- Modélisation du parc hydroélectrique et extraction des entrées du modèle mathématique de la base de données de l'unité production d'Hydro-Québec ;
- Choix d'une représentation des fonctions de production des centrales hydroélectriques qui réduit le plus possible l'écart entre la production estimée et la production réelle. En fait, les fonctions de production des centrales hydroélectriques ne sont ni linéaires, ni même concaves. Or, les modèles d'optimisation utilisés sont linéaires. D'où le besoin de linéariser ces fonctions tout en restant proche de leurs vraies valeurs. Plusieurs approches de linéarisation sont présentées et comparées entre elles ;
- Application et comparaison des méthodes SDDP et ASDP, avec le choix effectué pour la modélisation des fonctions de productions, sur la partie la plus importante du parc hydroélectrique du Québec, à savoir La Grande Rivière.

### 1.4 Contenu du mémoire

Le présent mémoire est divisé en sept chapitres, y compris cette introduction. Le chapitre 2 contient une revue de littérature des méthodes connues de résolution du problème de planification de la production d'hydroélectricité, en particulier les méthodes de programmation dynamique stochastique (SDP et SDDP). Le chapitre 3 présente plusieurs aspects de modélisation qui entrent en jeu dans le domaine de l'hydroélectricité en général, notamment les fonctions de production des centrales hydroélectriques et les apports hydriques naturels. Le

chapitre 4 décrit le problème de planification de la production d'hydroélectricité. Le chapitre 5 décrit les méthodes de programmation dynamique stochastique utilisées dans ce mémoire : SDP (en particulier ASDP) et SDDP. Le chapitre 6 présente et analyse les résultats numériques obtenus sur un cas d'étude réel : La Grande Rivière. Enfin, dans le dernier chapitre, on conclut sur les travaux qui ont été menés et on expose des pistes de recherche future à explorer.

## CHAPITRE 2 REVUE DE LITTÉRATURE

Ce chapitre présente une revue de littérature des méthodes utilisées dans la planification moyen-terme de la production d'hydroélectricité. La section 2.1 présente le problème de gestion de réservoirs sous une forme générale et classe les méthodes qui ont été utilisées pour le résoudre. Les sections 2.2, 2.3 et 2.4 traitent respectivement des méthodes basées sur les arbres de scénarios, des méthodes basées sur la programmation dynamique stochastique (SDP) et de la méthode de programmation dynamique duale stochastique (SDDP). Finalement, la section 2.5 est consacrée aux fonctions de production des centrales hydroélectriques qui constitue un élément central de modélisation dans le problème de planification de la production d'hydroélectricité.

### 2.1 Gestion de réservoirs sous incertitude

Si l'on fait l'hypothèse que les apports de chaque semaine, notamment  $\xi_1$  durant la première semaine, sont connus au début de celle-ci, le problème de planification moyen-terme de la production d'hydroélectricité s'écrit sous la forme suivante :

$$\begin{aligned} \Theta_1 = \max_{\substack{\mathbf{X}_1, \mathbf{S}_1 \\ \mathbf{S}_1 = f(\mathbf{S}_0, \xi_1, \mathbf{X}_1)}} B_1(\mathbf{S}_0, \xi_1, \mathbf{X}_1) + \mathbb{E}_{\xi_2} \left[ \max_{\substack{\mathbf{X}_2, \mathbf{S}_2 \\ \mathbf{S}_2 = f(\mathbf{S}_1, \xi_2, \mathbf{X}_2)}} B_2(\mathbf{S}_1, \xi_2, \mathbf{X}_2) + \right. \\ \left. \dots + \mathbb{E}_{\xi_t} \left[ \max_{\substack{\mathbf{X}_t, \mathbf{S}_t \\ \mathbf{S}_t = f(\mathbf{S}_{t-1}, \xi_t, \mathbf{X}_t)}} B_t(\mathbf{S}_{t-1}, \xi_t, \mathbf{X}_t) + \dots \right. \right. \\ \left. \left. \dots + \mathbb{E}_{\xi_T} \left[ \max_{\substack{\mathbf{X}_T, \mathbf{S}_T \\ \mathbf{S}_T = f(\mathbf{S}_{T-1}, \xi_T, \mathbf{X}_T)}} B_T(\mathbf{S}_{T-1}, \xi_T, \mathbf{X}_T) \right] \right] \right] \end{aligned} \quad (2.1)$$

où  $\mathbf{S}_0$  est le vecteur de l'état initial du système,  $\mathbf{S}_t$  est le vecteur de l'état du système à la fin de la période  $t$ ,  $\mathbf{X}_t$  est le vecteur des décisions de la période  $t$  (débits turbinés, débits déversés etc.),  $B_t$  est la fonction bénéfice immédiat de la période  $t$  et  $\xi_t$  vecteur des aléas ayant lieu durant la période  $t$ .

Le vecteur  $\mathbf{S}_t$  de l'état du système contient les volumes des réservoirs à la fin de la période  $t$ . Il peut aussi contenir d'autres composantes telles que des apports passés dans le cas où l'on modélise les apports hydriques naturels avec un modèle périodique auto-régressif PAR( $p$ ).

$\mathbf{S}_t = f(\mathbf{S}_{t-1}, \xi_t, \mathbf{X}_t)$  représente les contraintes de conservation de masse des réservoirs et des

contraintes sur les décisions de la période  $t$ . La fonction  $f$  peut être vue comme la fonction de transition du schéma de programmation dynamique du problème (2.1).

L'aléa  $\xi_t$  peut porter sur plusieurs paramètres : apports hydriques naturels au réservoirs, demande en électricité, prix de l'électricité etc. Dans cette revue, on se restreint au cas où l'aléa porte seulement sur les apports hydriques naturels.

Dans le contexte de planification moyen-terme, on ne tient pas compte des non-linéarités qui entrent en jeu dans les définitions exactes des fonctions  $f$  et  $B_t$ .

Le but ultime de la planification de la production d'hydroélectricité est de trouver une politique de gestion optimale sur tout l'horizon de planification. À la période  $t$ , à partir d'un état initial  $\mathbf{S}_{t-1}$  et un apport prévu  $\xi_t$ , il faut prendre une décision optimale  $\mathbf{X}_t$  qui maximise la somme des bénéfices immédiat  $B_t(\mathbf{S}_{t-1}, \xi_t, \mathbf{X}_t)$  et futur

$$\mathbb{E}_{\xi_{t+1}} \left[ \max_{\substack{\mathbf{X}_{t+1}, \mathbf{S}_{t+1} \\ \mathbf{S}_{t+1}=f(\mathbf{S}_t, \xi_{t+1}, \mathbf{X}_{t+1})}} B_{t+1}(\mathbf{S}_t, \xi_{t+1}, \mathbf{X}_{t+1}) + \dots + \mathbb{E}_{\xi_T} \left[ \max_{\substack{\mathbf{X}_T, \mathbf{S}_T \\ \mathbf{S}_T=f(\mathbf{S}_{T-1}, \xi_T, \mathbf{X}_T)}} B_T(\mathbf{S}_{T-1}, \xi_T, \mathbf{X}_T) \right] \right].$$

La politique de gestion doit permettre au décideur de prendre à chaque période une décision pesée qui représente un compromis entre les différentes réalisations de l'aléa futur. De manière schématique, si l'apport hydrique peut avoir deux réalisations possibles : humide (H) et sec (S), les décisions doivent prendre en compte les quatre scénarios représentés à la figure 2.1.

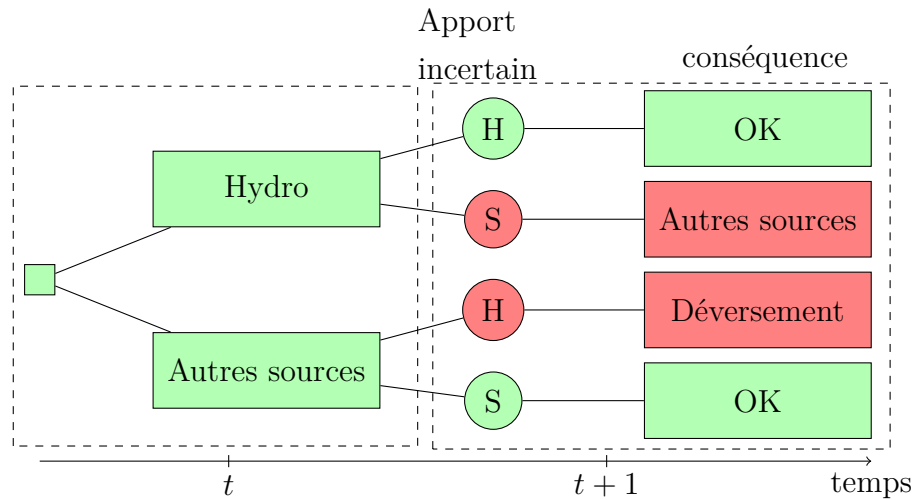


Figure 2.1 Processus décisionnel à deux étapes

Plusieurs méthodologies ont été proposées pour résoudre le problème (2.1). Des articles tels

que ceux de Yeh (1985), Labadie (2004) et plus récemment Rani et Moreira (2010), ont résumé ces méthodologies sous forme de revues qui ont grandement facilité le travail bibliographique des chercheurs dans ce domaine.

Il est à noter que le problème (2.1) est formulé avec l'hypothèse d'information parfaite de l'aléa  $\xi_t$  à une période  $t$  donnée alors que l'aléa futur  $(\xi_{t+1}, \xi_{t+2}, \dots, \xi_T)$  est stochastique. Cette hypothèse est réaliste dans le cadre de la planification moyen- et long-terme (voir, par exemple, Labadie (2004)).

Labadie (2004) classe les méthodes de résolution de (2.1) en deux catégories :

- **optimisation stochastique implicite** : dont la procédure est schématisée dans la figure 2.2. Dans ce type de méthodes, (2.1) n'est pas globalement résolu sur le processus stochastique  $(\xi_t)_t$ , mais plutôt sur un grand nombre  $M$  de scénarios possibles  $\xi_t^{(m)}$ ,  $t = 1 \dots T$ ,  $m = 1 \dots M$ . Chaque scénario est résolu via un modèle déterministe et a sa propre politique de gestion. Une politique de gestion globale est par la suite construite via des méthodes de régression à partir des politiques propres des  $M$  scénarios.

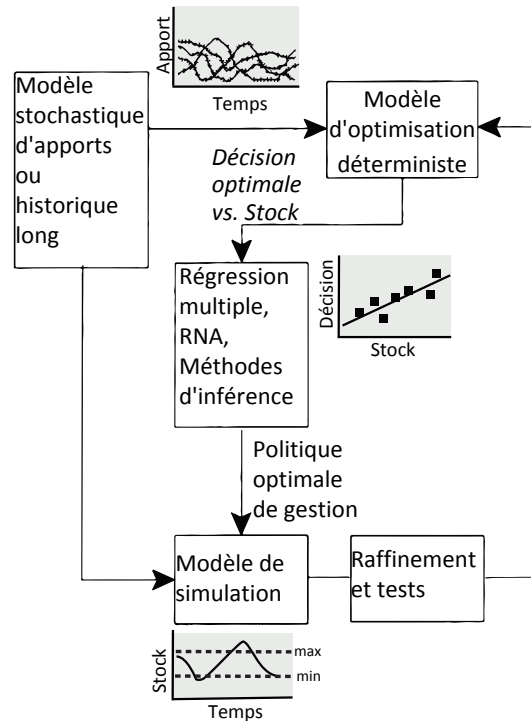


Figure 2.2 Procédure d'optimisation stochastique implicite (Labadie (2004))

L'article de Karamouz et Houck (1987) présente un exemple de ce type d'approche où il est question d'itérer un cycle constitué d'un programme dynamique déterministe, d'une analyse avec la régression et d'une simulation de politique.



L'inconvénient de ce type d'approches est que les politiques obtenues sont anticipatives. Cette anticipativité est due au fait que l'on résout des problèmes déterministes dont les décisions à une période  $t$  donnée ne prennent pas en compte toutes les réalisations possibles de l'aléa qui aura lieu dans les périodes subséquentes.

- **optimisation stochastique explicite** : dont la procédure est schématisée dans la figure 2.3. Dans ce cas, le processus stochastique  $(\xi_t)_t$  est explicitement pris en compte et les modèles utilisés sont stochastiques dans le sens que les décisions sont prises de sorte à maximiser la somme du bénéfice immédiat et du bénéfice futur. Ce bénéfice futur est explicitement formulé comme somme pondérée des bénéfices futurs de toutes les réalisations de l'aléa futur.

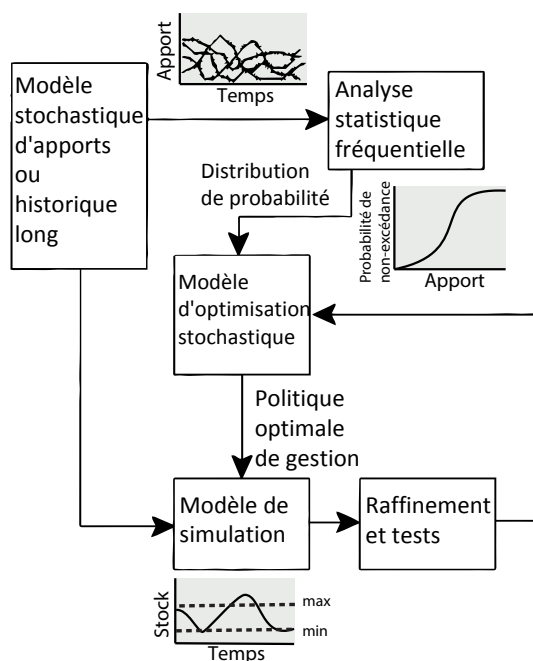


Figure 2.3 Procédure d'optimisation stochastique explicite (Labadie (2004))

Les politiques issues de ce type d'approches sont plus fiables vu qu'elles sont fidèles à la formulation rigoureuse du problème (2.1).

Dans la suite on expose trois types d'approches d'optimisation stochastique explicite, à savoir les méthodes basées sur les arbres de scénarios, les méthodes basées sur la programmation dynamique stochastique sous sa forme récursive pure (ASDP) et la méthode de programmation dynamique duale stochastique (SDDP) (Pereira et Pinto (1991)) qui est une variante itérative de ASDP basée sur la décomposition de Benders (Benders (1962)).

## 2.2 Résolution par réduction à un arbre de scénarios

Les méthodes basées sur une réduction à un arbre de scénarios du processus stochastique des apports partent d'une discrétisation de celui-ci sous forme d'un arbre de scénarios fini. Un arbre de scénarios est schématisé dans la figure 2.4.

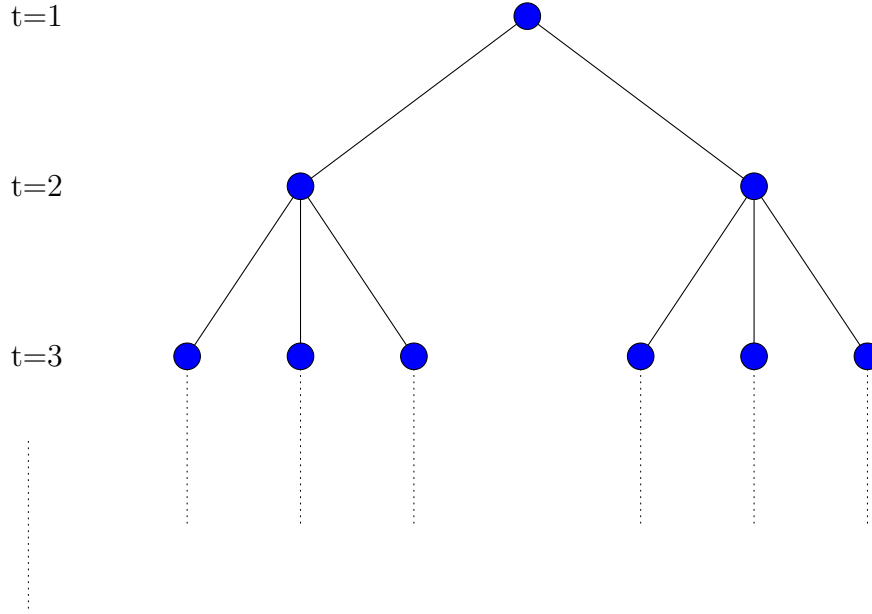


Figure 2.4 Un arbre de scénarios

Le problème (2.1) est reformulé de manière équivalente sur l'arbre de scénarios en associant à chaque nœud  $n$  de l'arbre un vecteur de décisions  $\mathbf{X}_n$  et en le liant à son nœud père avec des contraintes équivalentes aux contraintes de transition  $\mathbf{S}_t = f(\mathbf{S}_{t-1}, \boldsymbol{\xi}_t, \mathbf{X}_t)$  dans (2.1).

Le problème est ensuite résolu par une méthode de décomposition. Les méthodes de décomposition connues pour résoudre un programme stochastique sur arbre de scénarios sont de deux types : méthodes de décomposition par étape et méthodes de décomposition par scénario.

Les méthodes de décomposition par étape sont basées sur la décomposition de Benders (Benders (1962)) et affectent une fonction de bénéfice futur  $\mathcal{F}_n$  à chaque nœud  $n$  de l'arbre. Ces fonctions  $\mathcal{F}_n$  étant concaves linéaires par morceaux, elles sont construites de manière itérative par ajout d'hyperplans coupants. La méthode *L-Shaped* (Birge et Louveaux (2011)) est souvent utilisée dans ce contexte (Archibald et al. (1999)). Cette méthode peut être coûteuse dans le cas où l'arbre contient un grand nombre de nœuds. Pour contourner cette difficulté, Santos et Diniz (2009) proposent une approche multi-périodes agrégeant plusieurs périodes

consécutives en une seule pour diminuer le nombre de fonctions de bénéfice futur  $\mathcal{F}_n$  à définir.

Les méthodes de décomposition par scénario définissent un sous-problème pour chaque scénario. Le problème maître contient les contraintes liantes dites de non-anticipativité qui assurent qu'à chaque nœud, une seule décision est prise quelle que soit le scénario qui le traverse. La méthode la plus connue dans cette classe est le *Progressive Hedging*. Elle a été récemment implémentée en gestion de réservoirs par Carpentier et al. (2013).

Les méthodes basées sur la réduction à un arbre de scénarios restent limitées en matière de taille des arbres de scénarios. Ceci fait perdre au processus stochastique d'apport sa richesse et contraint sa discrétisation.

### 2.3 Résolution par programmation dynamique stochastique approchée

Le problème (2.1) peut être écrit sous la forme récursive équivalente suivante :

$$\begin{aligned} \Theta_1 &= \max_{\substack{\mathbf{X}_1, \mathbf{S}_1 \\ \mathbf{S}_1 = f(S_0, \boldsymbol{\xi}_1, \mathbf{X}_1)}} B_1(S_0, \boldsymbol{\xi}_1, \mathbf{X}_1) + \mathcal{F}_2(\mathbf{S}_1) \\ \mathcal{F}_t(\mathbf{S}_{t-1}) &= \mathbb{E}_{\boldsymbol{\xi}_t} \left[ \max_{\substack{\mathbf{X}_t, \mathbf{S}_t \\ \mathbf{S}_t = f(\mathbf{S}_{t-1}, \boldsymbol{\xi}_t, \mathbf{X}_t)}} B_t(\mathbf{S}_{t-1}, \boldsymbol{\xi}_t, \mathbf{X}_t) + \mathcal{F}_{t+1}(\mathbf{S}_t) \right] \quad t = 2 \dots T \\ \mathcal{F}_{T+1} &\equiv 0. \end{aligned} \tag{2.2}$$

En exploitant le principe d'optimalité de Bellman (Bellman (1957)), le problème (2.2) peut être résolu à l'optimalité. Les fonctions de bénéfice futur  $\mathcal{F}_t$  sont construites récursivement en allant de  $t = T$  à  $t = 2$ . La récurrence est bien définie puisque la fonction de bénéfice futur à la dernière période est soit connue soit nulle comme énoncé précédemment ( $\mathcal{F}_{T+1} \equiv 0$ ).

L'algorithme de SDP commence par construire une approximation  $\mathfrak{F}_T$  de  $\mathcal{F}_T$ .  $\mathfrak{F}_T$  est ensuite utilisée pour construire une approximation  $\mathfrak{F}_{T-1}$  de  $\mathcal{F}_{T-1}$  et ainsi de suite jusqu'à construire une approximation  $\mathfrak{F}_2$  de  $\mathcal{F}_2$ . À chaque période, l'approximation  $\mathfrak{F}_t$  et son gradient sont évalués sur une grille fine du domaine de définition de  $\mathbf{S}_{t-1}$ , ensuite des méthodes d'interpolation (Johnson et al. (1993), Tejada-Guibert et al. (1993)) ou d'extrapolation (programmation dynamique duale (DDP), Benders (1962)) sont utilisées pour construire une représentation complète de  $\mathfrak{F}_t$  en se basant sur la concavité de  $\mathcal{F}_t$  (Birge et Louveaux (2011)).

Malgré sa simplicité, la complexité de cette méthode est exponentielle en le nombre de réservoirs. En effet, supposons que le système comporte  $R$  réservoirs et que le volume de chaque

réservoir est discrétisé en  $M$  valeurs. Si l'état du système est représenté seulement par les volumes des réservoirs, la complexité de SDP est en  $\mathcal{O}(T \times M^R)$ .

Pour contourner sa complexité, deux approches ont été proposées.

La première, communément appelée méthode d'agrégation-désagrégation, consiste à agréger un ensemble de réservoirs en un seul puis lancer SDP sur le système agrégé. Une politique de gestion est ainsi obtenue sur le système agrégé et nécessite un post-traitement de désagrégation pour être opérationnelle. Ce type d'approches a été largement utilisé pour résoudre le problème de planification moyen- et long-terme dans un système multi-réservoirs afin de surmonter la combinatoire de SDP. Archibald et al. (1997) décomposent le problème en sous-problèmes dans chacun l'ensemble des réservoirs est divisé en trois réservoirs agrégés : les réservoirs en amont d'un réservoir  $r$ , le réservoir  $r$  et les réservoirs en aval de  $r$ . Turgeon et Charbonneau (1998) agrègent les réservoirs par rivière puis assemble toutes les rivières en un seul réservoir.

Malgré que l'approche agrégation-désagrégation surmonte la combinatoire du problème multi-réservoirs, des difficultés sont rencontrées dans le post-traitement de désagrégation (Saad et al. (1994)).

La deuxième approche utilisée avec SDP pour contourner sa combinatoire consiste à éviter d'évaluer la fonction de bénéfice futur sur tous les points d'une grille régulière mais plutôt sur une sous partie de cette grille. Cervellera et al. (2006) utilisent des techniques telles que les tableaux orthogonaux, les hypercubes latins et les suites à discrédance faible pour discrétiser efficacement l'espace d'état. Ayant évalué la fonction de bénéfice futur sur l'espace discrétisé, Cervellera et al. (2006) utilisent une technique basée sur les réseaux de neurones pour approximer cette fonction sur l'espace d'état entier. Plus récemment, Krau et al. (2015) ont développé une technique de discrétisation en temps réel de l'espace d'état en le découpant en hyper-rectangles.

## 2.4 Résolution par programmation dynamique duale stochastique

La méthode de programmation dynamique duale stochastique (SDDP) est une approche qui surmonte la complexité de SDP et n'assume aucune discrétisation *a priori* de l'espace d'état. Cette méthode a été présentée pour la première fois par Pereira et Pinto (1991) et a reçu une notoriété mondiale par la suite. Elle a été utilisée comme outil de gestion dans plusieurs systèmes hydroélectriques dans le monde : Brésil (Maceira et al. (2008)), Scandinavie (Røtting et al. (1992), Mo et al. (2001), Kristiansen (2004)), Nouvelle-Zélande (Halliburton et al. (2004)), Turquie (Tilmant et Kelman (2007)), Afrique (Goor et al. (2010a)),

Goor et al. (2010b), Tilmant et al. (2010)).

Au lieu de construire les approximations  $\mathfrak{F}_t$  de  $\mathcal{F}_t$  en une seule récurrence, SDDP procède de manière itérative et raffine à chaque itération une sur-approximation  $\mathfrak{F}_t$  de  $\mathcal{F}_t$  en exploitant la concavité de la fonction de bénéfice futur  $\mathcal{F}_t$  (Birge et Louveaux (2011)) avec des hyperplans coupants.

Plus précisément, une itération de SDDP est constituée d'une passe *vers l'avant* (*forward pass*) et d'une passe *vers l'arrière* (*backward pass*). La passe *vers l'avant* permet de préciser des candidats  $S_{t-1}$ ,  $t = 2 \dots T$  par simulation de la politique approximative  $(\mathfrak{F}_t)_t$  courante avec un (ou plusieurs) scénario(s) tiré(s) aléatoirement du processus stochastique d'apport. La passe *vers l'arrière* permet de raffiner la politique  $(\mathfrak{F}_t)_t$  par ajout d'une coupe de Benders à  $\mathfrak{F}_t$  au point candidat  $S_{t-1}$  pour  $t = T \dots 2$ .

Vu son caractère itératif, SDDP nécessite un critère d'arrêt. En plus du nombre maximum d'itérations, deux types de critères d'arrêt ont été présenté dans la littérature. Le premier (Pereira et Pinto (1991), Shapiro (2011)) consiste à calculer une borne supérieure en résolvant le problème à une étape de la première période. Un estimateur de la borne inférieure est ensuite calculé en simulant un nombre assez grand de scénarios du processus stochastique d'apport et permet de définir un intervalle de confiance sur la valeur de l'objectif (2.1). Le critère consiste à tester si la borne supérieure est à l'intérieur de cet intervalle de confiance. Ce critère est discuté par Shapiro (2011), en particulier plus la tolérance de l'intervalle de confiance est petite plus celui-ci est large et SDDP s'arrête prématurément. Le deuxième critère d'arrêt est présenté par Homem-de Mello et al. (2011) sous forme de tests d'hypothèses.

Sous certaines hypothèses non restrictives, Philpott et Guan (2008) et Shapiro (2011) ont prouvé la convergence de SDDP après un nombre fini d'itérations. La preuve faite par Philpott et Guan (2008) est plus simple et se base sur la preuve de convergence de la méthode de décomposition de Benders (Benders (1962)). Le résultat central utilisé dans cette preuve est la finitude du nombre de coupes de Benders pouvant être générées pour définir les fonctions de bénéfice futur approximées  $(\mathfrak{F}_t)_t$ . Rigoureusement parlant, les deux articles précités prouvent que  $\mathfrak{F}_t$  converge presque sûrement vers la vraie fonction de bénéfice futur  $\mathcal{F}_t$  pour tout  $t = 2 \dots T$  et ce après un nombre fini d'itérations.

Le seul inconvénient de SDDP réside dans la phase de *rodage* dont elle a besoin avant de stabiliser la politique. En effet, les coupes générées aux premières itérations ne sont plus significatives après un certain nombre d'itérations. En plus, plus le nombre de périodes de l'horizon est grand plus cette phase de *rodage* est longue. Quelques articles ont proposé des approches visant à accélérer la convergence de SDDP et réduire l'impact de la phase de *rodage* sur son temps d'exécution.

de Matos et al. (2015) propose deux techniques pour accélérer SDDP. La première consiste à ne pas simuler un nombre fixe de scénarios dans la phase *vers l'avant* mais plutôt d'augmenter ce nombre progressivement au fil des itérations. La deuxième technique consiste à sélectionner quelques coupes définissant les fonctions de bénéfice futur  $\mathfrak{F}_t$  lorsque l'on résout les problèmes à une étape. Ces deux techniques réussissent à améliorer le temps d'exécution de SDDP. Cependant, la deuxième peut dégrader la qualité de la politique de gestion comme le souligne de Matos et al. (2015).

Aouam et Yu (2008) propose une autre technique pour améliorer la performance de SDDP qui consiste à ajouter, au début de l'algorithme, des coupes de Benders exclusivement aux fonctions de bénéfice futur  $\mathfrak{F}_t$  pour  $t = 2 \dots T'$  où  $T' < T$  et ensuite exécuter l'algorithme standard de SDDP sur tout l'horizon de planification.

## 2.5 Fonctions de production des centrales hydroélectriques

La puissance électrique  $p$  (MW) générée par une centrale hydroélectrique s'exprime en fonction de la hauteur de chute  $h$  (m) et le débit turbiné  $q$  ( $\text{m}^3/\text{s}$ ) :

$$p(q, h) = \eta(q) \rho g q h \quad (2.3)$$

où  $\rho$  ( $\text{kg}/\text{m}^3$ ) est la densité volumique de l'eau,  $g$  ( $\text{m}/\text{s}^2$ ) est l'accélération de la pesanteur.

En pratique, la centrale ne capte pas toute l'énergie potentielle de l'eau à cause des pertes dans les conduites communes et l'effet de tarage aval.

Dans les publications de Hammadia (2000), Breton et al. (2002) et Breton et al. (2004), ces pertes ont été prises en compte dans la hauteur de chute de manière sous-jacente en remplaçant  $h$  dans (2.3) par la hauteur de chute nette  $h_{nette}$  telle que  $h_{nette} = h - \Delta h_{pertes}$ .

Dans l'article de Diniz et Maceira (2008), ces pertes sont prises en compte de manière explicite. Une représentations quadri-dimensionnelle est adoptée  $(q, h, \Delta h_{pertes}) \mapsto p$  où les pertes énergétiques sont prises en compte dans la hauteur de chute ( $\Delta h_{pertes}$ ) mais en tant que variable explicite.

Pour les besoins de la planification moyen-terme de la production hydroélectrique, une fonction globale du domaine *débit-hauteur de chute* au domaine *puissance* doit être construite au préalable. Dans le cas où l'on utilise des modèles linéaires qui maximise la production, cette fonction globale doit être approximée par des fonctions concaves linéarisables.

Plusieurs articles traitant la planification de la production d'hydroélectricité (Archibald et al. (1999), Tilmant et Kelman (2007)) négligent l'effet hauteur de chute dans leurs modèles.

Cependant cette hypothèse simplificatrice reste acceptable pour des centrales dont la hauteur de chute fluctue peu autour d'une valeur nominale.

La procédure d'approximation concave la plus utilisée pour atteindre ce but consiste à approximer cette fonction par une fonction concave linéaire par morceaux en utilisant des méthodes d'ajustement convexe linéaire par morceaux comme celle développée par Magnani et Boyd (2009). Ce type d'approche a été utilisé, entre autres, par Goor et al. (2010a).

## CHAPITRE 3 ÉLÉMENTS DE MODÉLISATION

Ce chapitre présente les éléments de modélisation dans le contexte de l'hydroélectricité. En premier lieu, les infrastructures et installations en jeu dans ce domaine sont introduites. Ensuite, plusieurs aspects importants de modélisation sont étudiés en détail, à savoir les fonctions de production des centrales hydroélectriques, les apports naturels (en eau) aux réservoirs et la modélisation de la fluctuation de la demande en électricité par des classes.

**Note :** Les images réelles présentes dans ce chapitre sont prises du site-web <http://www.hydroquebec.ca>.

### 3.1 Infrastructures et installations hydroélectriques

L'hydroélectricité est produite par l'intermédiaire d'une double conversion énergétique représentée dans la figure 3.1 :

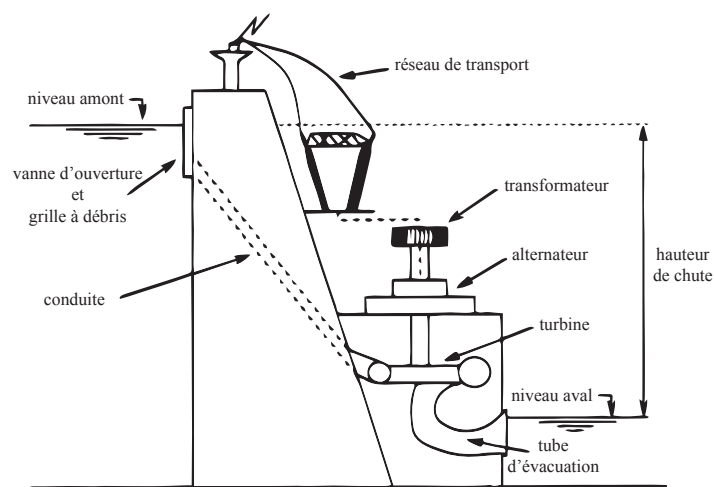


Figure 3.1 Processus de production de l'hydroélectricité (Hammadia (2000))

– *Conversion de l'énergie potentielle de pesanteur en énergie mécanique*

À l'ouverture de la vanne, l'eau emmagasinée dans le réservoir en amont s'engouffre dans la conduite amenant à une turbine. La vitesse de l'eau entraîne la rotation et de la turbine et du rotor de l'alternateur. Ensuite

– *Conversion de l'énergie mécanique de rotation en énergie électrique*

L'alternateur formé du rotor (mobile) et du stator (fixe) produit un courant alternatif grâce à la rotation du rotor.



Quelles sont alors les caractéristiques des infrastructures et installations utilisées pour produire l'hydroélectricité ?

### 3.1.1 Réservoir

Un réservoir est une infrastructure faisant barrage à l'eau qui coule le long d'une rivière, il permet ainsi de stocker l'énergie potentielle de l'eau. La figure 3.2 nous donne une image d'un réservoir réel.



Figure 3.2 Barrage de la Trenché (Québec)

Un réservoir (schéma de base dans la figure 3.3) est caractérisé par :

- un volume minimum  $v_{min}$  et un volume maximum  $v_{max}$ . Le volume  $v_{max} - v_{min}$  est le volume pour générer de l'électricité. Ainsi, dans les modèles qu'on verra par la suite, on considère le volume *relatif* (à  $v_{min}$ ) du réservoir dont les bornes sont 0 et  $\bar{v} = v_{max} - v_{min}$  ;
- une fonction  $v \mapsto Niv$  donnant le niveau de l'eau du réservoir en fonction de son volume. Cette fonction est concave linéaire par morceaux (on peut imaginer la surface horizontale de l'eau augmenter avec le niveau).

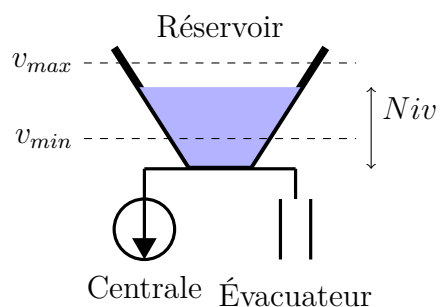


Figure 3.3 Schéma de base d'un réservoir

### 3.1.2 Centrale hydroélectrique

Une centrale hydroélectrique est une usine de production d'électricité constituée de plusieurs groupes turbines-alternateurs. Ces derniers peuvent être parallèlement exploités en totalité ou en partie. Les groupes turbines-alternateurs, et donc les centrales hydroélectriques, sont caractérisés par une fonction entrée-sortie appelée fonction *débit-puissance* qui dépend de la hauteur de chute. Cette fonction sera étudiée dans la section 3.2.

On distingue deux types de centrales :

- *centrale à réservoir* : installée en aval d'un réservoir. Ce type de centrales exploitent souvent une dénivellation de l'eau importante et produisent une grande puissance. La figure 3.4 donne une image de ce type de centrales.

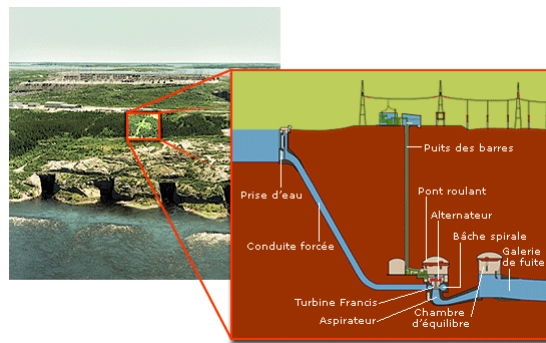


Figure 3.4 Centrale Robert-Bourassa, une centrale souterraine à réservoir (Québec)

- *centrale au fil de l'eau* : alimentée directement par un cours d'eau et ne disposant pas de réservoir. Ce type de centrales ont une faible hauteur de chute et par conséquent la puissance qu'elles produisent dépend uniquement du débit du cours d'eau. La figure 3.5 donne une image de ce type de centrales.

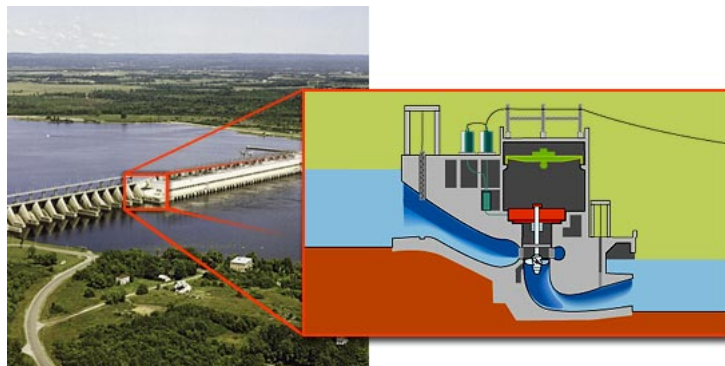


Figure 3.5 Centrale de Carillon, une centrale au fil de l'eau (Québec)

### 3.1.3 Ouvrage d'évacuation

Un ouvrage d'évacuation (appelé aussi *évacuateur de crues* ou *déversoir*) est un aménagement hydroélectrique en aval d'un réservoir qui permet de contrôler son niveau d'eau pour empêcher qu'il ne dépasse sa limite. Cet ouvrage permet aussi d'approvisionner en eau un réservoir plus bas. La figure 3.6 donne une image d'un évacuateur de crues réel.

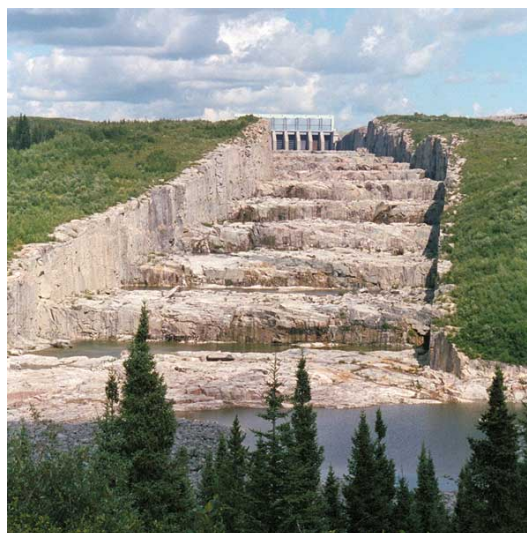


Figure 3.6 Évacuateur de crues du barrage Robert-Bourassa (Québec)

## 3.2 Fonction de production d'une centrale hydroélectrique

Le fonction de production d'une centrale hydroélectrique dépend de deux facteurs : le débit turbiné et la hauteur de chute. Cette dernière étant la différence entre le niveau en amont de la centrale et le niveau en aval.

L'énergie potentielle de l'eau que la centrale capte et transforme effectivement en énergie électrique est plus petite que l'énergie potentielle de l'eau en repos (avant d'être acheminée et turbinée). En effet, plusieurs facteurs affectent cette énergie captée, à savoir

- les frottements au niveaux des conduites d'eau et d'autres installations d'acheminement
- le tarage aval qui correspond à l'augmentation du niveau aval due au débit sortant de la centrale.

On modélise souvent cette perte énergétique par une perte en hauteur de chute. On parle ainsi de *hauteur de chute brute* (celle de l'eau en repos), notée  $h$ , et de *hauteur de chute nette*

(celle ressentie par la centrale), notée  $h_{nette}$ . La relation entre ces deux hauteurs de chute est

$$h_{nette} = h - \Delta h_{pertes} \quad (3.1)$$

où  $\Delta h_{pertes}$  correspond aux pertes énergétiques.

### 3.2.1 Cas d'un groupe turbine-alternateur

La puissance électrique  $p_{T/A}$  (MW) produite par un groupe turbine-alternateur opérant à un débit  $q$  ( $m^3/s$ ) et une hauteur de chute brute  $h$  (m) s'écrit sous la forme :

$$p_{T/A}(q, h) = \eta_{T/A}(q) \rho g q h_{nette} \quad (3.2)$$

où  $\rho$  ( $kg/m^3$ ) est la densité volumique de l'eau,  $g$  ( $m/s^2$ ) l'accélération de la pesanteur et  $\eta_{T/A}(q) < 1$  le rendement du groupe. Ce rendement dépend du débit turbiné  $q$  comme le montre la figure 3.7, en particulier il est optimal pour un débit spécifique  $q^{tracé}$  appelé *débit de tracé*.

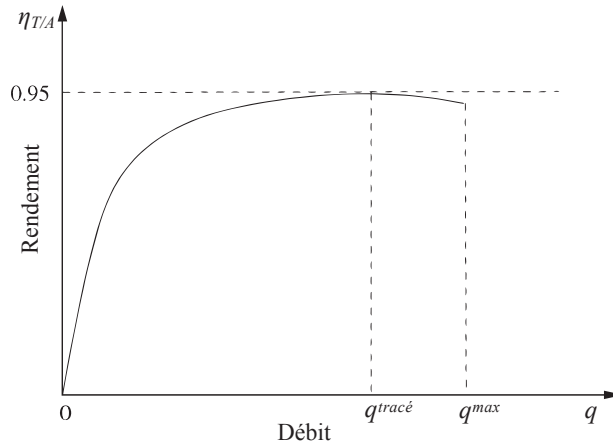


Figure 3.7 Rendement d'un groupe turbine-alternateur (Hammadia (2000))

On dispose souvent de la courbe de la fonction débit-puissance d'un groupe turbine-alternateur pour une hauteur de chute brute spécifique appelée hauteur de référence, notée  $h^0$ . On note cette fonction  $q \mapsto p_{T/A}^0(q)$ . La figure 3.8 donne un exemple de fonction débit-puissance

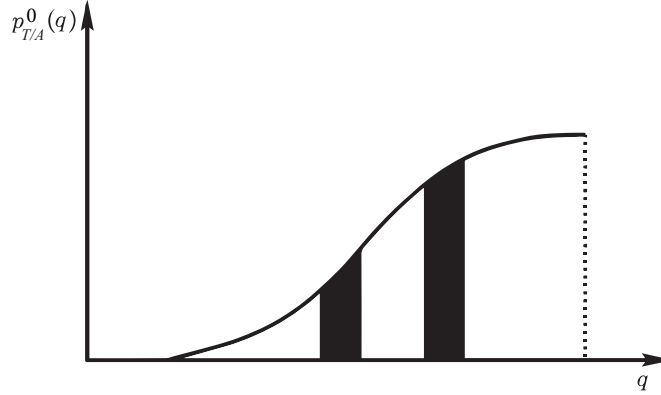


Figure 3.8 Exemple de fonction débit-puissance (Hammadia (2000))

Les zones noircies sont des zones d'opération interdites (phénomènes de vibration ou cavitation).

On peut aussi obtenir la valeur de la puissance à un débit  $q$  et hauteur de chute brute  $h$  donnés par la relation de similitude suivante (Hammadia (2000), Breton et al. (2004)) :

$$p_{T/A}(q, h) = \left( \frac{h_{nette}}{h_{nette}^0} \right)^\beta p_{T/A}^0 \left( \left( \frac{h_{nette}}{h_{nette}^0} \right)^{-\alpha} q \right) \quad (3.3)$$

où  $\alpha$  et  $\beta \in \mathbb{R}$  sont des paramètres qui dépendent du groupe turbine-alternateur. (3.3) peut être approximée par

$$p_{T/A}(q, h) = \left( \frac{h}{h^0} \right)^\beta p_{T/A}^0 \left( \left( \frac{h}{h^0} \right)^{-\alpha} q \right) \quad (3.4)$$

Cette approximation repose sur les deux hypothèses suivantes :

- $\frac{h_{nette}}{h_{nette}^0} = \frac{h - \Delta h_{pertes}}{h^0 - \Delta h_{pertes}^0} \approx \frac{h}{h^0}$  le rapport des hauteurs de chute nettes est sensiblement égal à celui des hauteurs de chute brutes ;
- $\eta_{T/A} \left( \left( \frac{h_{nette}}{h_{nette}^0} \right)^{-\alpha} q \right) \approx \eta_{T/A} \left( \left( \frac{h}{h^0} \right)^{-\alpha} q \right) \approx \eta_{T/A}(q)$ . En pratique, cette hypothèse n'est pas restrictive car d'une part le groupe turbine-alternateur fonctionne généralement à son débit de tracé (sur la figure 3.7, on peut observer le caractère aplati de la courbe  $q \mapsto \eta_{T/A}(q)$  autour du débit de tracé), et d'une autre part  $\left( \frac{h_{nette}}{h_{nette}^0} \right)^{-\alpha} \approx \left( \frac{h}{h^0} \right)^{-\alpha} \approx 1$ .

La figure 3.9 illustre cette transformation de la fonction débit-puissance de la hauteur de chute brute de référence à une autre hauteur de chute brute. Dans cette figure,  $\alpha(h)$  représente

$\left(\frac{h}{h^0}\right)^{-\alpha}$  et  $\beta(h)$  représente  $\left(\frac{h}{h^0}\right)^{\beta}$ .

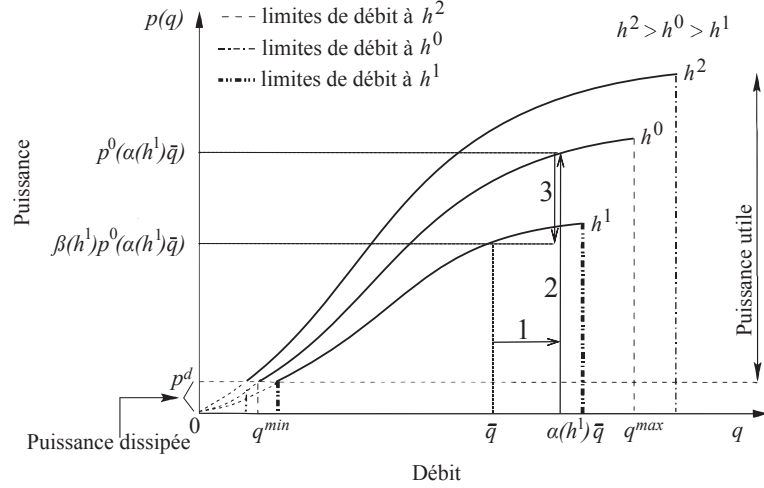


Figure 3.9 Transformation par similitude de la fonction débit-puissance (Hammadia (2000))

### 3.2.2 Cas d'une centrale

Une centrale hydroélectrique est composée d'un ou plusieurs groupes turbines-alternateurs. La figure 3.10 donne un exemple de configuration physique d'une centrale.

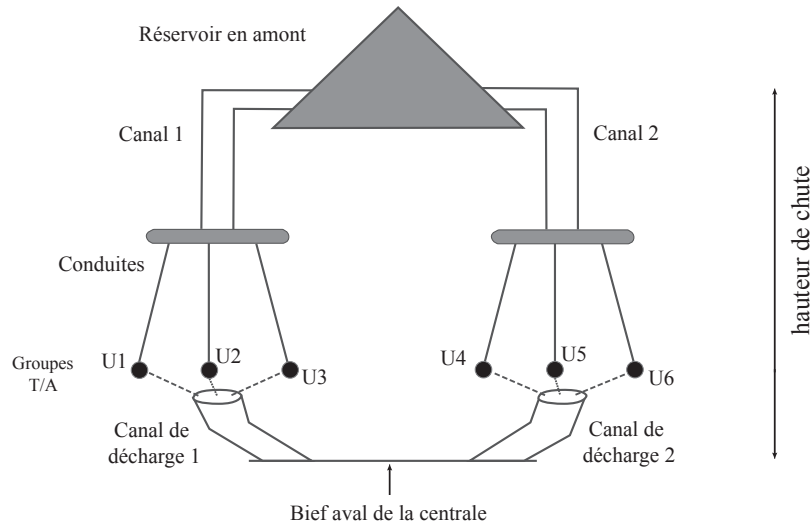


Figure 3.10 Exemple de configuration physique d'une centrale (Hammadia (2000))

La puissance générée par la centrale pour un débit turbiné total  $q$ , une hauteur de chute nette  $h$  et un nombre  $U$  de groupes turbines-alternateurs engagés s'exprime comme la somme des puissances générées par ses groupes turbines-alternateurs

$$p_C(q, h) = \sum_{i=1}^U p_{T/A}(\delta_i q, h) \quad (3.5)$$

où les variables (positives)  $\delta_i$  vérifient  $\sum_{i=1}^U \delta_i = 1$ , le débit total  $q$  est reparti sur les groupes turbines-alternateurs de façon à maximiser la puissance totale générée par la centrale tout en respectant les contraintes de zones interdites des groupes turbines-alternateurs (Problème de chargement optimal connu dans la littérature (e.g. Hammadia (2000), Breton et al. (2004))).

La formule (3.3) de transformation par similitude pour un groupe turbine-alternateur peut être généralisée au cas d'une centrale

$$\begin{aligned} p_C(q, h) &= \sum_{i=1}^U p_{T/A}(\delta_i q, h) \\ &= \sum_{i=1}^U \left( \frac{h}{h^0} \right)^\beta p_{T/A}^0 \left( \left( \frac{h}{h^0} \right)^{-\alpha} \delta_i q \right) \\ &= \left( \frac{h}{h^0} \right)^\beta \sum_{i=1}^U p_{T/A}^0 \left( \delta_i \left( \frac{h}{h^0} \right)^{-\alpha} q \right) \\ &= \left( \frac{h}{h^0} \right)^\beta p_C^0 \left( \left( \frac{h}{h^0} \right)^{-\alpha} q \right), \end{aligned} \quad (3.6)$$

où la fonction  $p_C^0$  est la fonction de production de la centrale à la hauteur de chute de référence  $h^0$ .

### 3.2.3 Modélisation pour la planification moyen-terme de la production

Dans le cadre de la planification moyen-terme de la production hydroélectrique (horizon d'une à deux années avec un pas de temps hebdomadaire), il est question de déterminer la quantité de ressources hydrauliques à puiser de chaque réservoir pour répondre à la demande et de savoir la puissance électrique obtenue.

La planification court-terme (horizon de quelques jours avec un pas horaire) tient compte des décisions globales prises par la planification moyen-terme et fournit un plan d'opération des centrales avec toutes les considérations liées au choix des groupes turbines-alternateurs à engager ou retirer pour la maintenance, les zones d'opérations interdites des groupes etc.

Dans cette perspective, la fonction de production d'une centrale hydroélectrique pour un modèle moyen-terme est modélisée par un ensemble fini de configurations de production. Plus précisément, pour une centrale  $C$  composée de  $U_C$  groupes turbines-alternateurs, on dispose de points  $(q_i^0, p_i^0)$ ,  $i \in \{0, \dots, U_C\} \cup \{U_C + 1, \dots, U_C + L_C\}$  correspondant à des configurations optimales de production à la hauteur de chute (brute) de référence. Les points  $(q_i^0, p_i^0)$ ,  $i \in \{0, \dots, U_C\}$  correspondent à l'engagement de  $i$  groupes turbines-alternateurs. Les points  $(q_i^0, p_i^0)$ ,  $i \in \{U_C + 1, \dots, U_C + L_C\}$  correspondent à des configurations qui engagent tous les  $U_C$  groupes avec un débit supplémentaire, auxquelles on peut avoir recours en cas de demande élevée. Pour alléger la notation, on note dorénavant  $U'_C = U_C + L_C$ .

Pour éviter d'utiliser des modèles d'optimisation en nombres entiers pour la planification moyen-terme, on modélise la fonction de production à hauteur de chute de référence par interpolation des points  $(q_i^0, p_i^0)$ ,  $i \in \{0, \dots, U'_C\}$ , on obtient ainsi une fonction concave (voir, par exemple, Hammadia (2000)) linéaire par morceaux. La figure 3.11 donne un exemple d'une telle interpolation.

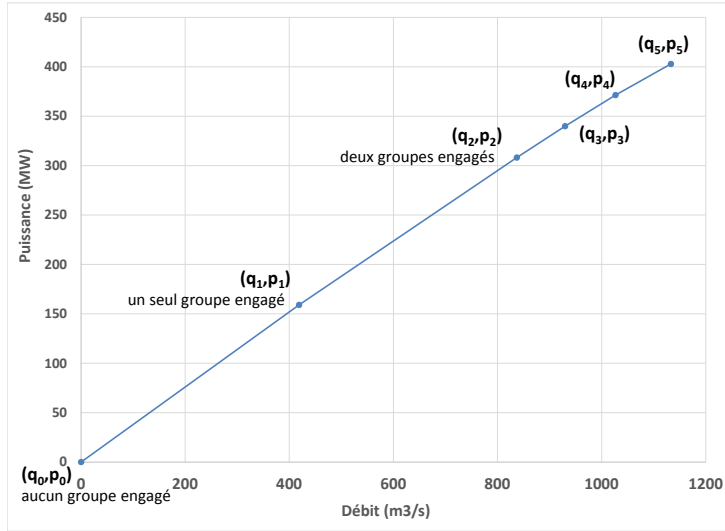


Figure 3.11 Exemple de fonction de production à hauteur de référence (Centrale à deux groupes)

Si, par exemple, le modèle que l'on optimise indique qu'il faut turbiner au pas  $t$  de durée  $\Delta_t$  à un débit  $q$  tel que  $q_1^0 < q < q_2^0$ , on interprète cela comme turbiner pendant  $\lambda \Delta_t$  au débit  $q_1^0$  et pendant  $(1 - \lambda) \Delta_t$  au débit  $q_2^0$ , où  $\lambda \in [0, 1]$  vérifie  $q = \lambda q_1^0 + (1 - \lambda) q_2^0$ .



À une autre hauteur de chute brute  $h^1$ , on peut obtenir les points  $(q_i^1, p_i^1)$ ,  $i \in \{0, \dots, U'_C\}$  correspondants aux mêmes configurations de production via la formule de transformation par similitude (3.6), i.e.

$$\begin{cases} q_i^1 &= \left(\frac{h^1}{h^0}\right)^\alpha q_i^0 \\ p_i^1 &= \left(\frac{h^1}{h^0}\right)^\beta p_i^0 \end{cases} \quad (3.7)$$

Ainsi, on peut obtenir une fonction de production pour tout l'espace  $(q, h)$ . La figure 3.12 en donne un exemple.

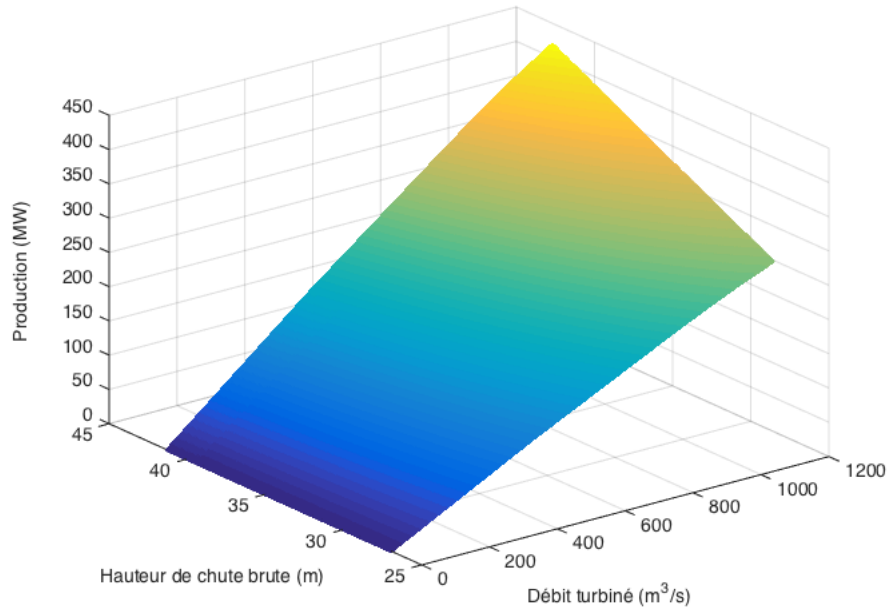


Figure 3.12 Exemple de fonction de production globale

Cette fonction, quoique concave si l'on fixe la hauteur de chute, n'est pas globalement concave. Ainsi pour l'utiliser dans un cadre général de programmation linéaire où l'on maximise la production hydroélectrique, il faut l'approximer par une fonction concave au préalable. Des méthodes pour l'approximer par une fonction concave sont introduites dans la section suivante.

### 3.2.4 Approximation concave des fonctions de production

Plusieurs approches d'approximation concave des fonctions de production sont envisageables. On présente ici deux méthodes différentes :

- approximation par enveloppe concave de deux fonctions de production aux hauteurs de chute limites (minimum et maximum) ;
- approximation par une fonction concave affine par morceaux.

#### Approximation par enveloppe concave

On obtient les fonctions de production respectives de la centrale à ses hauteurs de chute brutes minimum  $h^{min}$  et maximum  $h^{max}$  via les équations (3.7). Ce qui nous donne les points  $(q_i^{min}, p_i^{min})$   $(q_i^{max}, p_i^{max})$ ,  $i \in \{0, \dots, U'_C\}$

La production  $p_C(q, h)$  est alors approximée par

$$\begin{aligned}
 p_C(q, h) &\approx \tilde{p}_C(q, h) := \max_{\lambda, \mathbf{p}} \mathbf{p} \\
 \text{s. à } q &= \sum_{i=1}^{U'_C} \lambda_i^{min} q_i^{min} + \sum_{i=1}^{U'_C} \lambda_i^{max} q_i^{max} \\
 h &= \sum_{i=1}^{U'_C} \lambda_i^{min} h^{min} + \sum_{i=1}^{U'_C} \lambda_i^{max} h^{max} \\
 \mathbf{p} &= \sum_{i=1}^{U'_C} \lambda_i^{min} p_i^{min} + \sum_{i=1}^{U'_C} \lambda_i^{max} p_i^{max} \\
 \sum_{i=1}^{U'_C} \lambda_i^{min} + \sum_{i=1}^{U'_C} \lambda_i^{max} &= 1 \\
 \lambda &\geq 0.
 \end{aligned} \tag{3.8}$$

Dorénavant, on fait référence à cette approche par **EnvConc**.

#### Approximation par une fonction concave affine par morceaux

La fonction de production globale  $p_C(q, h)$  est approximée par une fonction concave affine par morceaux, i.e.

$$\begin{aligned}
 p_C(q, h) &\approx \tilde{p}_C(q, h) := \min_{j=1 \dots K} (a_{q,j}q + a_{h,j}h + b_j) \\
 &= \max_{\mathbf{p}} \mathbf{p} \\
 \mathbf{p} &\leq a_{q,j}q + a_{h,j}h + b_j \quad j = 1 \dots K.
 \end{aligned} \tag{3.9}$$

Ici,

- $K$  : nombre d'hyperplans affines qui approximent la fonction  $p_C(q, h)$ , c'est un paramètre ;
- $(a_{q,j}, a_{h,j}, b_j)$ ,  $j = 1 \dots K$  sont des variables à déterminer.

Pour déterminer les valeurs de ces variables, on utilise un algorithme présenté par Magnani et Boyd (2009) qui essaie d'ajuster un nuage de points  $u_i, y_i \in \mathbb{R}^n \times \mathbb{R}$  par une fonction convexe linéaire par morceaux. Dans notre cas, le nuage de points consiste en  $(q_i, h_i), p_i \in \mathbb{R}^2 \times \mathbb{R}$  qu'on essaie d'ajuster avec une fonction concave linéaire par morceaux.

À partir des points  $(q_i^0, p_i^0)$ ,  $i \in \{0, \dots, U'_C\}$  qui correspondent aux fonctionnements de la centrale à la hauteur de chute de référence, on obtient via (3.7) les points  $(q_i^j, p_i^j)$ ,  $i \in \{0, \dots, U'_C\}$  pour différentes hauteurs de chute  $h^j$ ,  $j = 1 \dots N_h$ . On obtient ainsi le nuage de points  $(q_i^j, h^j), p_i^j$   $(i, j) \in \{0, \dots, U'_C\} \times \{1, \dots, N_h\}$ . Sans restreindre la généralité et pour alléger la description de l'algorithme de Magnani et Boyd (2009), on note ce nuage de points  $(q_i, h_i), p_i$   $i = 1 \dots N_h \times U'_C$ .

L'algorithme part d'une partition  $P_1^{(0)}, \dots, P_K^{(0)}$  initiale de l'ensemble des indices des éléments du nuage de points  $\{1 \dots N_h \times U'_C\}$ , trouve un hyperplan optimal au sens d'un critère donné pour chaque ensemble de la partition, met à jour la partition et refait la même procédure jusqu'à ce que la partition ne change pas ou qu'un nombre maximum d'itérations est atteint.

---

### ALGORITHME DE PARTITIONNEMENT

---

- **Données** : partition  $P_1^{(0)}, \dots, P_K^{(0)}$  de  $\{1 \dots N_h \times U'_C\}$  et nombre maximum d'itérations  $l_{max}$
  - **Pour**  $l = 0 \dots l_{max}$ 
    1. **Pour**  $j = 1 \dots K$   
trouver un hyperplan affine  $(a_{q,j}^{(l+1)}, a_{h,j}^{(l+1)}, b_j^{(l+1)})$  approximant le nuage de points restreint à l'ensemble  $P_j^{(l)}$
    2. Construire une nouvelle partition  $P_1^{(l+1)} \dots P_K^{(l+1)}$  avec les nouveaux hyperplans, i.e.  

$$P_j^{(l+1)} = \left\{ i, j = \min \left\{ r, \min_{s=1 \dots K} \left( a_{q,s}^{(l+1)} q_i + a_{h,s}^{(l+1)} h_i + b_s^{(l+1)} \right) = a_{q,r}^{(l+1)} q_i + a_{h,r}^{(l+1)} h_i + b_r^{(l+1)} \right\} \right\}$$
    3. **STOP** si  $P_j^{(l+1)} = P_j^{(l)}$  pour tout  $j = 1 \dots K$
- 

Le point **2.** de l'algorithme ci-dessus est, en d'autres termes, l'affectation à la nouvelle  $j^e$  partition  $P_j^{(l+1)}$  les points où le  $j^e$  hyperplan affine  $(a_{q,j}, a_{h,j}, b_j)$  est *actif*.

Le point **1.** de l'algorithme ci-dessus peut avoir plusieurs variantes.

- **Meilleur hyperplan affine au sens des moindres carrés** : cette variante est celle proposée dans l'article de Magnani et Boyd (2009) consiste à déterminer  $(a_{q,j}, a_{h,j}, b_j)$  avec le critère de moindres carrés, i.e.

$$\mathbf{X}_j^{(l+1)*} = (a_{q,j}^{(l+1)}, a_{h,j}^{(l+1)}, b_j^{(l+1)}) \in \underset{(a_{q,j}, a_{h,j}, b_j)}{\operatorname{argmin}} \left( \sum_{i \in P_j^{(l)}} (a_{q,j} q_i + a_{h,j} h_i + b_j - p_i)^2 \right). \quad (3.10)$$

Ce qui équivaut à résoudre

$$\underbrace{\begin{pmatrix} \sum_{i \in P_j^{(l)}} q_i^2 & \sum_{i \in P_j^{(l)}} q_i h_i & \sum_{i \in P_j^{(l)}} q_i \\ \sum_{i \in P_j^{(l)}} q_i h_i & \sum_{i \in P_j^{(l)}} h_i^2 & \sum_{i \in P_j^{(l)}} h_i \\ \sum_{i \in P_j^{(l)}} q_i & \sum_{i \in P_j^{(l)}} h_i & |P_j^{(l)}| \end{pmatrix}}_{A_j^{(l+1)}} \underbrace{\begin{pmatrix} a_{q,j}^{(l+1)} \\ a_{h,j}^{(l+1)} \\ b_j^{(l+1)} \end{pmatrix}}_{\mathbf{X}_j^{(l+1)}} = \underbrace{\begin{pmatrix} \sum_{i \in P_j^{(l)}} q_i p_i \\ \sum_{i \in P_j^{(l)}} h_i p_i \\ \sum_{i \in P_j^{(l)}} p_i \end{pmatrix}}_{w_j^{(l+1)}}. \quad (3.11)$$

Dans le cas où la matrice  $A_j^{(l+1)}$  est inversible, i.e. le problème (3.10) admet une unique solution optimale  $\mathbf{X}_j^{(l+1)*}$  telle que

$$\mathbf{X}_j^{(l+1)*} = (A_j^{(l+1)})^{-1} w_j^{(l+1)}. \quad (3.12)$$

Dans le cas contraire, i.e. le problème (3.10) admet une infinité de solution, plusieurs options peuvent être suivies. L'une d'elles est de trouver la solution  $\mathbf{X}_j^{(l+1)*}$  la plus proche de la solution de l'itération précédente  $X_j^{(l)}$  au sens de la norme euclidienne

$$\mathbf{X}_j^{(l+1)*} \in \underset{\mathbf{X}_j^{(l+1)}}{\operatorname{argmin}} \left\{ \|\mathbf{X}_j^{(l+1)} - \mathbf{X}_j^{(l)*}\|_2, \quad A_j^{(l+1)} \mathbf{X}_j^{(l+1)} = w_j^{(l+1)} \right\}. \quad (3.13)$$

Dorénavant, on fait référence à cette variante par **PartMC**.

- **Meilleur hyperplan affine au sens du minimum du maximum des écarts** : ici, on propose de remplacer le critère de moindres carrés par le critère de minimisation du maximum des écarts, i.e.

$$\mathbf{X}_j^{(l+1)*} = (a_{q,j}^{(l+1)}, a_{h,j}^{(l+1)}, b_j^{(l+1)}) \in \underset{(a_{q,j}, a_{h,j}, b_j)}{\operatorname{argmin}} \left( \max_{i \in P_j^{(l)}} |a_{q,j} q_i + a_{h,j} h_i + b_j - p_i| \right). \quad (3.14)$$

Ce qui revient à résoudre le programme linéaire suivant

$$\begin{aligned}
& \min_{z, a_{q,j}, a_{h,j}, b_j} && z \\
& \text{s. à} && z \geq a_{q,j}q_i + a_{h,j}h_i + b_j - p_i \\
& && z \geq -(a_{q,j}q_i + a_{h,j}h_i + b_j - p_i).
\end{aligned} \tag{3.15}$$

Cette variante peut être coûteuse lorsque les programmes linéaires (3.15) sont de taille relativement grande.

Dorénavant, on fait référence à cette variante par **PartMINMAX**.

**Remarque :** Au cours de l'algorithme, il se peut qu'un des ensembles de la partition devient vide. Dans ce cas on diminue le nombre d'hyperplans.

Comme le mentionne Magnani et Boyd (2009), l'algorithme de partitionnement qu'on vient de développer dépend fortement de la partition de départ  $P_1^{(0)}, \dots, P_K^{(0)}$ . C'est pourquoi il propose de faire plusieurs essais avec différentes partitions initiales et d'en garder la meilleure approximation par hyperplans obtenue.

---

#### ALGORITHME DE PARTITIONNEMENT++

---

- **Données :** Nombre d'essais  $N_{essais}$  et nombre maximum d'itérations  $l_{max}$
  - **Pour**  $s = 1 \dots N_{essais}$ 
    1. Générer une partition initiale aléatoire
    2. Exécuter ALGORITHME DE PARTITIONNEMENT avec la partition générée et  $l_{max}$
    3. Garder une trace de la meilleure approximation par hyperplans obtenue
- 

Le point **1.** de l'algorithme ci-dessus est fait, comme le suggère Magnani et Boyd (2009), de la manière suivante

- $K$  points deux à deux différents  $u_j = (q_j, h_j)$ ,  $j = 1 \dots K$  du nuages sont tirés au hasard
- $P_j^0 = \{i, \|(q_i, h_i) - u_j\|_2 < \|(q_i, h_i) - u_s\|_2 \ \forall s \neq j\}$ ,  $j = 1 \dots K$ . Autrement dit  $P_j^0$  est formé des points qui sont plus proches de  $u_j$  que des autres  $u_s$ ,  $s \neq j$

Le point **3.** de l'algorithme ci-dessus dépend du critère que l'on utilise dans l'algorithme de partitionnement.

Si l'on est dans le cas de **PartMC**, le meilleur partitionnement est celui qui a la plus petite erreur-type (RMSE) (en anglais « Root Mean Square Error »), qui s'écrit

$$RMSE = \sqrt{\frac{1}{N_h \times U'_C} \sum_{i=1}^{N_h \times U'_C} \left( \left( \min_{j=1 \dots K} a_{q,j} q_i + a_{h,j} h_i + b_j \right) - p_i \right)^2}. \quad (3.16)$$

Si l'on est dans le cas de **PartMINMAX**, le meilleur partitionnement est celui qui a la plus petite erreur maximum (ErreurMax), qui s'écrit

$$ErreurMax = \max_{i=1 \dots N_h \times U'_C} \left| \left( \min_{j=1 \dots K} a_{q,j} q_i + a_{h,j} h_i + b_j \right) - p_i \right|. \quad (3.17)$$

Dans le chapitre 6, une comparaisons des trois approches (**EnvConc**, **PartMC**, **PartMINMAX**) est faite sur plusieurs centrales hydroélectriques. On montre que **PartMC** et **PartMINMAX** sont plus appropriées aux fonctions de productions des centrales considérées avec un léger avantage pour **PartMINMAX** grâce à sa simplicité.

### 3.3 Apports hydriques naturels

#### 3.3.1 Modèles auto-régressifs périodiques

Les apports hydriques naturels sont généralement caractérisés par des aspects périodiques (dus aux saisons de l'année) et stochastiques. La figure 3.13 illustre ces deux aspects en traçant la courbe des apports par période hebdomadaire d'un historique de 53 années (la période 1 correspond à la 1<sup>re</sup> semaine de janvier et ainsi de suite jusqu'à la 52<sup>e</sup> période qui correspond à la dernière semaine de décembre).

Si l'on considère le cas d'une seule rivière, l'apport hydrique durant une période  $t$  donnée peut être modélisé par une seule variable aléatoire  $\mathbf{A}_t$ . La corrélation temporelle entre les apports de différentes périodes est souvent modélisée par un modèle PAR( $p$ ) (Periodic Autoregressive model of order  $p$ ) dans lequel l'apport  $\mathbf{A}_t$  est fonction des apports passés  $\mathbf{A}_{t-1}, \dots, \mathbf{A}_{t-p}$  par la formule

$$\frac{\mathbf{A}_t - \mu_t}{\sigma_t} = \sum_{j=1}^p \phi_{t,j} \frac{\mathbf{A}_{t-j} - \mu_{t-j}}{\sigma_{t-j}} + \epsilon_t. \quad (3.18)$$

Où  $\mu_i$  et  $\sigma_i$  sont respectivement l'espérance mathématique et l'écart-type de  $\mathbf{A}_i$ ,  $\phi_{t,1}, \dots, \phi_{t,p}$  sont les paramètres du modèle PAR( $p$ ) et  $\epsilon_t$  un bruit gaussien de moyenne nulle et d'écart-type  $\sigma_{\epsilon_t}$ . Les bruits  $\epsilon_t$ ,  $t = 1 \dots T$  sont indépendants. Il est à noter que la définition (3.18) est incomplète sans la donnée des  $p$  apports hydriques passés (quantités déterministes connues)

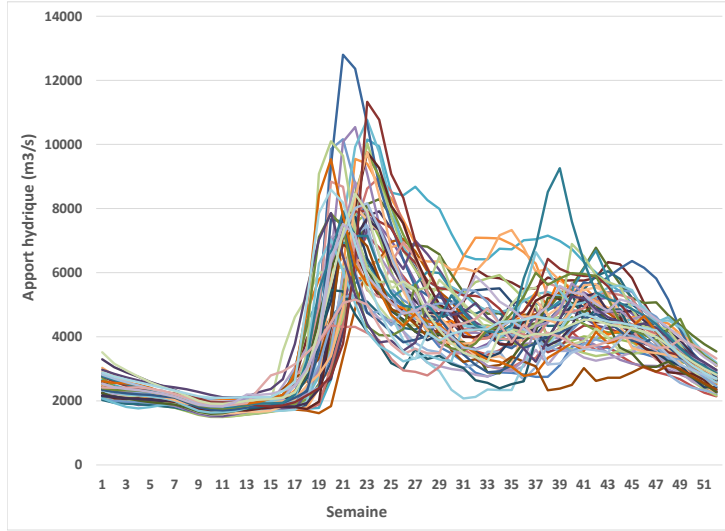


Figure 3.13 Historique des apports hydriques (53 années)

$A_0, \dots, A_{1-p}$ , où  $A_{1-i}$  est l'apport hydriques à la rivière durant la  $i^{\text{e}}$  période avant le début de l'horizon considéré et ce en remontant dans le temps.

Dans le cadre de la programmation dynamique stochastique, plus le paramètre  $p$  est grand plus la complexité de la résolution augmente (exponentiellement). C'est pour cette raison que souvent on utilise juste un PAR(1). L'équation (3.18) devient

$$\frac{\mathbf{A}_t - \mu_t}{\sigma_t} = \phi_t \frac{\mathbf{A}_{t-1} - \mu_{t-1}}{\sigma_{t-1}} + \epsilon_t. \quad (3.19)$$

L'équation (3.19) contient une incohérence. En effet, le bruit  $\epsilon_t$  étant gaussien, il peut prendre toute valeur réelle, même négative, ce qui peut emmener à des réalisations de signe négatif de l'apport  $\mathbf{A}_t$ . Pour éviter cela, une option serait d'utiliser plutôt une loi gaussienne tronquée pour le bruit  $\epsilon_t$ . Une solution plus générale et qui présente aucune incohérence est d'utiliser une transformation  $f$  de l'apport qui n'est pas affectée par les valeurs négatives que peut prendre le bruit  $\epsilon_t$ , le modèle PAR(1) prend ainsi la forme

$$\frac{f(\mathbf{A}_t) - \mu_t}{\sigma_t} = \phi_t \frac{f(\mathbf{A}_{t-1}) - \mu_{t-1}}{\sigma_{t-1}} + \epsilon_t. \quad (3.20)$$

Un des candidats potentiels pour jouer le rôle de la fonction  $f$  dans (3.20) est la fonction

logarithme népérien, notée  $\ln$ , on obtient alors le modèle suivant

$$\frac{\ln(\mathbf{A}_t) - \mu_t}{\sigma_t} = \phi_t \frac{\ln(\mathbf{A}_{t-1}) - \mu_{t-1}}{\sigma_{t-1}} + \epsilon_t. \quad (3.21)$$

C'est ce modèle que l'on utilise dorénavant.

### 3.3.2 Estimation des paramètres du modèle PAR(1)

Les paramètres du modèle PAR(1) transformé (3.21) sont estimés à partir de l'historique des apports. On dispose souvent d'un nombre  $H$  d'années de données historiques  $A_t^{(i)}$ ,  $t = 1 \dots T$ ,  $i = 1 \dots H$ . On peut ainsi faire les estimations suivantes, dans cet ordre et pour tout  $t = 1 \dots T$  :

1.  $\mu_t \approx \hat{\mu}_t = \frac{1}{H} \sum_{i=1}^H \ln(A_t^{(i)})$  ;
2.  $\sigma_t \approx \hat{\sigma}_t = \sqrt{\frac{1}{H} \sum_{i=1}^H (\ln(A_t^{(i)}) - \hat{\mu}_t)^2}$  ;
3.  $\phi_t$  est obtenu par régression linéaire au sens des moindres carrés appliquée à (3.21), i.e.

$$\phi_t \in \underset{\varphi}{\operatorname{argmin}} \sum_{i=1}^H \left( \frac{\ln(A_t^{(i)}) - \hat{\mu}_t}{\hat{\sigma}_t} - \varphi \frac{\ln(A_{t-1}^{(i)}) - \hat{\mu}_{t-1}}{\hat{\sigma}_{t-1}} \right)^2 ;$$

$$4. \sigma_{\epsilon_t} \approx \hat{\sigma}_{\epsilon_t} = \sqrt{\frac{1}{H} \sum_{i=1}^H \left( \frac{\ln(A_t^{(i)}) - \hat{\mu}_t}{\hat{\sigma}_t} - \phi_t \frac{\ln(A_{t-1}^{(i)}) - \hat{\mu}_{t-1}}{\hat{\sigma}_{t-1}} \right)^2}.$$

### 3.3.3 Génération d'un processus stochastique PAR(1) discrétisé

Dans le cadre de l'optimisation stochastique approchée, on approxime souvent les vraies distributions (continues) des aléas par des distributions discrètes. Cette démarche est largement utilisée en optimisation sur arbres de scénarios, en programmation dynamique stochastique etc.

Les modèles PAR subissent aussi cette démarche classique et peuvent être discrétisé sous forme d'un arbre de scénarios comme suit :

à partir de la valeur déterministe  $A_0$ , générer  $N$  réalisations de la variable aléatoire  $\mathbf{A}_1$  via l'équation (3.21). À partir de chaque réalisation de  $\mathbf{A}_1$  générer  $N$  réalisations de la variable aléatoire  $\mathbf{A}_2$ . Et ainsi de suite jusqu'à atteindre la période  $T$ . La figure 3.14 illustre cette démarche pour  $N = 2$ .



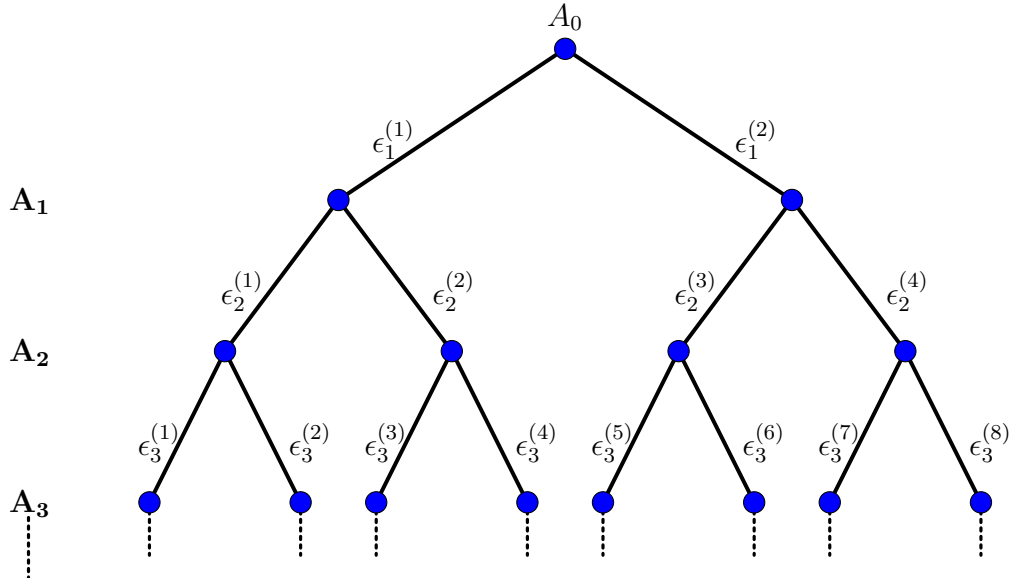


Figure 3.14 Génération naïve du processus stochastique discrétisé

En programmation dynamique stochastique et dans le cas où le processus des apports est régi par un modèle PAR(1), la fonction de valeur de Bellman de la période  $t$  dépend de l'apport passé  $\mathbf{A}_{t-1}$ . Comme on va le voir dans le chapitre 5, quand on utilise la transformation (équation (3.21)) des apports, cette fonction de Bellman n'est plus convexe par rapport à  $\mathbf{A}_{t-1}$ . Dans ce cas-ci, on est obligé de définir une fonction de valeur de Bellman propre à chaque valeur de l'apport passé  $\mathbf{A}_{t-1}$ . Or, on peut observer dans la figure 3.14 que le nombre de valeurs discrétisées de  $\mathbf{A}_{t-1}$  croît exponentiellement avec  $t$  ( $2^{t-1}$  si  $N = 2$ ) ce qui rend inefficace cette démarche de discrétisation du processus stochastique des apports.

Une alternative plus efficace pour discrétiser le processus stochastique des apports est d'échantillonner *a priori* la variable aléatoire d'apport  $\mathbf{A}_t$  en  $N$  quantiles  $\tilde{A}_t^{(i)}$ ,  $i = 1 \dots N$  pour toute période  $t = 1 \dots T$ . On passe alors d'une variable aléatoire continue  $\mathbf{A}_t$  à une variable aléatoire discrète  $\tilde{\mathbf{A}}_t$ . Cette discrétisation *a priori* se fait directement à partir de l'historique sans utiliser les variables aléatoires  $\epsilon_t$ . Ces dernières sont toutefois cruciales pour approximer les matrices des probabilités de transition  $P^t$  où  $P_{ij}^t = \mathbb{P}(\tilde{\mathbf{A}}_t = \tilde{A}_t^{(j)} | \tilde{\mathbf{A}}_{t-1} = \tilde{A}_{t-1}^{(i)})$ .

On obtient ainsi un processus stochastique PAR(1) des apports hydriques naturels sous forme d'une chaîne de Markov à temps discret dont l'espace des états est fini.

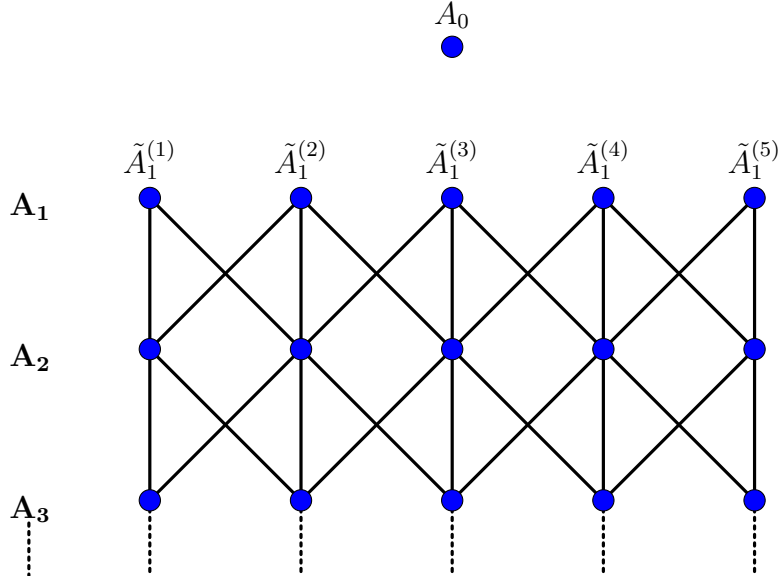


Figure 3.15 Génération efficace du processus stochastique discrétisé

### Estimation des matrices $P^t$ des probabilités de transition

Pour chaque période  $t = 1 \dots T$ ,  $i = 1 \dots N$ ,  $\tilde{A}_t^{(i)}$  représente le quantile  $\frac{i-1}{N-1}$  de l'historique de  $\mathbf{A}_t$

Soient  $t \in \{1, \dots, T\}$ ,  $i \in \{1, \dots, N\}$  et soient  $\tilde{\epsilon}_t^{(i,1)}, \tilde{\epsilon}_t^{(i,2)}, \dots, \tilde{\epsilon}_t^{(i,N)}$  tels que

$$\tilde{\epsilon}_t^{(i,N)} = \frac{\log(\tilde{A}_t^{(j)}) - \mu_t}{\sigma_t} - \phi_t \frac{\log(\tilde{A}_{t-1}^{(i)}) - \mu_{t-1}}{\sigma_{t-1}} \quad j = 1 \dots N. \quad (3.22)$$

Puisque  $\tilde{A}_t^{(1)} < \tilde{A}_t^{(2)} < \dots < \tilde{A}_t^{(N)}$ ,  $\tilde{\epsilon}_t^{(i,1)} < \tilde{\epsilon}_t^{(i,2)} < \dots < \tilde{\epsilon}_t^{(i,N)}$

Les probabilités de transition de  $\tilde{A}_{t-1}^{(i)}$  à  $\tilde{A}_t^{(j)}$  sont définies par

$$P_{ij}^t \approx \begin{cases} \mathbb{P}(\epsilon_t \in (-\infty, \epsilon_t^{(i,1)}]) & \text{si } j = 1 \\ \mathbb{P}(\epsilon_t \in (\epsilon_t^{(i,j-1)}, \epsilon_t^{(i,j)}]) & \text{si } 1 < j < N \\ \mathbb{P}(\epsilon_t \in (\epsilon_t^{(i,N-1)}, +\infty)) & \text{si } j = N \end{cases} \quad (3.23)$$

$$= \begin{cases} \frac{1}{\sigma_{\epsilon_t} \sqrt{2\pi}} \int_{-\infty}^{\epsilon_t^{(i,1)}} e^{-\frac{1}{2} \frac{y^2}{\sigma_{\epsilon_t}^2}} dy & \text{si } j = 1 \\ \frac{1}{\sigma_{\epsilon_t} \sqrt{2\pi}} \int_{\epsilon_t^{(i,j-1)}}^{\epsilon_t^{(i,j)}} e^{-\frac{1}{2} \frac{y^2}{\sigma_{\epsilon_t}^2}} dy & \text{si } 1 < j < N \\ \frac{1}{\sigma_{\epsilon_t} \sqrt{2\pi}} \int_{\epsilon_t^{(i,N-1)}}^{+\infty} e^{-\frac{1}{2} \frac{y^2}{\sigma_{\epsilon_t}^2}} dy & \text{si } j = N \end{cases}$$

### 3.4 Classes de demande

Durant une journée, la demande en électricité (puissance) peut varier. On distingue souvent des heures de pointe, au matin par exemple, et des heures de consommation moyenne. Il est alors judicieux d'inclure plusieurs classes de demande dans le problème de planification pour que l'estimation des ressources utilisées soit plus fidèle à la fluctuation de la demande.

Chaque classe de demande  $cls$  est représentée par une proportion de durée  $\delta_{cls}$  par rapport à la durée totale  $\Delta_t$  du pas de temps  $t$ . Par exemple, pour un pas hebdomadaire (i.e. 168 heures) la classe de pointe a une proportion de 0.25, soit une durée de 42 heures, et la classe normale 0.75, soit une durée de 126 heures. La figure 3.16 illustre la différence de demande en puissance pour un cas de deux classes

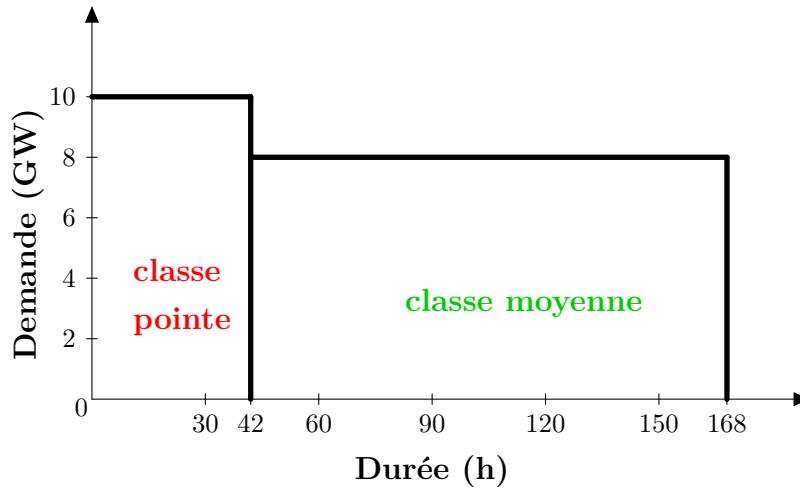


Figure 3.16 Classes de demande

Le modèle de planification contient des variables de décisions (débit turbiné, puissance générée) pour chaque classe de demande.

## CHAPITRE 4 PROBLÈME DE PLANIFICATION DE LA PRODUCTION D'HYDROÉLECTRICITÉ

Ce chapitre présente la notation et la formulation du problème de planification de la production d'hydroélectricité.

### 4.1 Notation

#### Symboles et ensembles

Symbole	Description
$\text{hm}^3$	hectomètres cube (un million de mètres cube)
MW	mega-watts
MWh	mega-watts heure
\$	dollar canadien
$\mathcal{R}$	ensemble des réservoirs
$\mathcal{R}^+$	ensemble des réservoirs à capacité positive
$\mathcal{R}^-$	ensemble des réservoirs sans capacité
$\mathcal{C}$	ensemble des centrales hydroélectriques
$\mathcal{C}^+$	ensemble des centrales à hauteur de chute variable (i.e., centrales à réservoir)
$\mathcal{C}^-$	ensemble des centrales à hauteur de chute fixe (i.e., centrales au fil de l'eau)
$\mathcal{OE}$	ensemble des ouvrages d'évacuation
$\mathcal{C}^d(r)$	ensemble des centrales du réservoir $r$ ( $d$ pour <i>downstream</i> )
$\mathcal{OE}^d(r)$	ensemble des ouvrages d'évacuation du réservoir $r$
$\mathcal{C}^u(r)$	ensemble des centrales dont l'eau turbinée va au réservoir $r$ ( $u$ pour <i>upstream</i> )
$\mathcal{OE}^u(r)$	ensemble des ouvrages d'évacuation dont l'eau déversée va au réservoir $r$
$r$	réservoir
$c$	centrale
$oe$	ouvrage d'évacuation
$r(c)$	réservoir de la centrale $c$
$l_r$	fonction donnant le niveau d'eau du réservoir $r$ en fonction de son volume
$p_c$	fonction donnant la puissance générée de la centrale $c$ en fonction de son débit (cas d'une centrale au fil de l'eau)
$B_t$	fonction de bénéfice immédiat de la semaine $t$

## Constantes

Symbole	Unité	Description
$\overline{v}_r$	hm <sup>3</sup>	volume maximum du réservoir $r$
$v_{r,0}$	hm <sup>3</sup>	volume initial du réservoir $r$
$V_0$	hm <sup>3</sup>	vecteur des volumes initiaux des réservoirs, $V_0 = (v_{r,0})_{r \in \mathcal{R}^+}$
$\delta_{r,t}$		proportion du réservoir $r$ par rapport à l'apport hydrique total à la rivière durant la semaine $t$
$l_c^d$	m	niveau aval de référence de la centrale $c$
$h_c^0$	m	hauteur de chute de référence de la centrale $c$
$\overline{q}_c^0$	m <sup>3</sup> /s	débit maximum de la centrale $c$ à sa hauteur de chute de référence
$\overline{q}_{oe}$	m <sup>3</sup> /s	débit maximum de l'ouvrage d'évacuation $oe$
$\gamma_c, \omega_c$		paramètres de la fonction de production de la centrale $c$
$\mathcal{K}_c$		nombre d'hyperplans affines approximant la fonction de production de la centrale $c$
$a_{c,q}^{(j)}$	MW/m <sup>3</sup> /s	pente relative au débit du j <sup>e</sup> hyperplan affine
$a_{c,h}^{(j)}$	MW/m	pente relative à la hauteur de chute du j <sup>e</sup> hyperplan affine
$b_c^{(j)}$	MW	ordonnée à l'origine du j <sup>e</sup> hyperplan affine
$\Delta$	h	durée d'une semaine ( $\Delta = 168\text{h}$ )
$\tau^p$		proportion de la classe de pointe par rapport à la durée d'une semaine
$\tau^{op}$		proportion de la classe moyenne par rapport à la durée d'une semaine $t$ ( $\tau^{op} = 1 - \tau^p$ )
$\rho$		constante de conversion d'un débit constant (m <sup>3</sup> /s) durant une semaine à un volume (hm <sup>3</sup> ), $\rho = \frac{3600 \times \Delta}{10^6} = 0.6048$
$d_t^p$	MW	demande minimum en puissance instantanée durant la semaine $t$ et la classe de demande de pointe
$d_t^{op}$	MW	demande minimum en puissance instantanée durant la semaine $t$ et la classe de demande moyenne
$\nu_t^p$	\$/MWh	prix de l'électricité durant la semaine $t$ et la classe de pointe
$\nu_t^{op}$	\$/MWh	prix de l'électricité durant la semaine $t$ et la classe moyenne
$c_t^p$	\$/MWh	coût encouru pour une demande non satisfaite durant la semaine $t$ et la classe de pointe
$c_t^{op}$	\$/MWh	coût encouru pour une demande non satisfaite durant la semaine $t$ et la classe moyenne

$T$	nombre de semaines dans l'horizon de planification (52 semaines pour une année, 104 pour 2 années ...)
-----	--

### Variables

Symbole	Unité	Description
$\mathbf{v}_{r,t}$	hm <sup>3</sup>	volume du réservoir $r$ à la fin de la semaine $t$
$\mathbf{l}_{r,t}^u$	m	niveau d'eau du réservoir $r$ durant la semaine $t$
$\mathbf{q}_{oe,t}$	m <sup>3</sup> /s	débit déversé de l'ouvrage d'évacuation $oe$ durant la semaine $t$
$\mathbf{h}_{c,t}$	m	hauteur de chute de la centrale $c$ durant la semaine $t$
$\mathbf{q}_{c,t}^p$	m <sup>3</sup> /s	débit turbiné de la centrale $c$ durant la semaine $t$ et la classe de demande de pointe ( $p$ pour <i>peak</i> )
$\mathbf{q}_{c,t}^{op}$	m <sup>3</sup> /s	débit turbiné de la centrale $c$ durant la semaine $t$ et la classe de demande moyenne ( $op$ pour <i>off-peak</i> )
$\mathbf{p}_{c,t}^p$	MW	puissance instantanée générée par la centrale $c$ durant la semaine $t$ et la classe de demande de pointe
$\mathbf{p}_{c,t}^{op}$	MW	puissance instantanée générée par la centrale $c$ durant la semaine $t$ et la classe de demande moyenne
$\mathbf{p}_t^p$	MW	puissance instantanée totale (somme sur toutes les centrales) générée durant la semaine $t$ et la classe de demande de pointe
$\mathbf{p}_t^{op}$	MW	puissance instantanée totale (somme sur toutes les centrales) générée durant la semaine $t$ et la classe de demande moyenne
$\mathbf{sh}_t^p$	MW	demande instantanée non satisfaite durant la semaine $t$ et la classe de demande de pointe ( $sh$ pour <i>shedding</i> )
$\mathbf{sh}_t^{op}$	MW	demande instantanée non satisfaite durant la semaine $t$ et la classe de demande moyenne
$\mathbf{V}_t$	hm <sup>3</sup>	vecteur des volumes des réservoirs à <b>fin</b> de la semaine $t$ , $\mathbf{V}_t = (\mathbf{v}_{r,t})_{r \in \mathcal{R}^+}$
$\mathbf{X}_t$	m <sup>3</sup> /s	vecteur des décisions de la semaine $t$ , $\mathbf{X}_t = ((\mathbf{q}_{c,t}^p)_{c \in \mathcal{C}}, (\mathbf{q}_{c,t}^{op})_{c \in \mathcal{C}}, (\mathbf{q}_{oe,t})_{oe \in \mathcal{OE}})$

### Processus stochastique d'apports

Symbole	Unité	Description
$A_1$	m <sup>3</sup> /s	apport hydrique à la rivière durant la semaine 1 (connu à l'avance)
$\mathbf{A}_t$	m <sup>3</sup> /s	variable aléatoire (continue) de l'apport à la rivière durant la semaine $t$

## 4.2 Formulation du problème

Si seuls les apports hydriques sont stochastiques et s'ils suivent un modèle PAR(1), le problème de planification moyen-terme de la production d'hydroélectricité peut se formuler comme suit

$$\begin{aligned} \Theta_1 = \max_{\substack{\mathbf{X}_1, \mathbf{V}_1 \\ \mathbf{V}_1 = f(V_0, A_1, \mathbf{X}_1)}} B_1(V_0, A_1, \mathbf{X}_1) + \mathbb{E}_{\mathbf{A}_2 | A_1} \left[ \max_{\substack{\mathbf{X}_2, \mathbf{V}_2 \\ \mathbf{V}_2 = f(\mathbf{V}_1, \mathbf{A}_2, \mathbf{X}_2)}} B_2(\mathbf{V}_1, \mathbf{A}_2, \mathbf{X}_2) + \right. \\ \left. \dots + \mathbb{E}_{\mathbf{A}_t | \mathbf{A}_{t-1}} \left[ \max_{\substack{\mathbf{X}_t, \mathbf{V}_t \\ \mathbf{V}_t = f(\mathbf{V}_{t-1}, \mathbf{A}_t, \mathbf{X}_t)}} B_t(\mathbf{V}_{t-1}, \mathbf{A}_t, \mathbf{X}_t) + \dots \right. \right. \\ \left. \left. \dots + \mathbb{E}_{\mathbf{A}_T | \mathbf{A}_{T-1}} \left[ \max_{\substack{\mathbf{X}_T, \mathbf{V}_T \\ \mathbf{V}_T = f(\mathbf{V}_{T-1}, \mathbf{A}_T, \mathbf{X}_T)}} B_T(\mathbf{V}_{T-1}, \mathbf{A}_T, \mathbf{X}_T) \right] \right] \right] \end{aligned} \quad (4.1)$$

où

- $B_1(V_0, A_1, \mathbf{X}_1)$  est la fonction bénéfice de la semaine 1, partant de l'état initial  $V_0$ , et ayant l'apport hydrique  $A_1$  et ayant pris les décisions  $\mathbf{X}_1$  ;
- Pour  $t = 2 \dots T$ ,  $B_t(\mathbf{V}_{t-1}, \mathbf{A}_t, \mathbf{X}_t)$  est la fonction bénéfice de la semaine  $t$ , partant de l'état initial  $\mathbf{V}_{t-1}$  (qui est en fait l'état final de la semaine  $t - 1$ ), et ayant un apport hydrique  $\mathbf{A}_t$  et ayant pris les décisions  $\mathbf{X}_t$  ;
- $f(\cdot, \cdot, \cdot)$  est une fonction de transition qui, partant d'un état initial au début d'une semaine et un apport hydrique et des décisions durant cette semaine, renvoie l'état final du système.

### Hypothèse :

Dans la formulation ci-haut, on a fait l'hypothèse d'*information parfaite* ("Wait and see"), i.e., l'apport durant une période est connu d'avance et influence directement les décisions prises durant cette période. En toute rigueur, cette hypothèse n'est pas indispensable. Cependant, elle permet de rendre le problème plus réaliste vu que :

- les périodes de l'horizon de planification dans un contexte moyen-terme sont souvent des semaines (parfois des mois). Il n'est donc pas réaliste de décider sans tenir compte de l'apport hydrique durant une période donnée ;
- des modèles météorologiques et statistiques permettent au décideur d'avoir des prévisions assez fiables à court-terme.

### 4.3 Description du problème

#### 4.3.1 Objectifs

Pour la semaine 1, partant de l'état initial du système  $V_0$  et sachant qu'il y aura l'apport hydrique  $A_1$ , l'objectif est de prendre une décision  $\mathbf{X}_1$  de sorte à maximiser la somme du bénéfice immédiat

$$B_1(V_0, A_1, \mathbf{X}_1) = \nu_1^p \tau^p \Delta \mathbf{p}_1^p + \nu_1^{op} \tau^{op} \Delta \mathbf{p}_1^{op} - c_1^p \tau^p \Delta \mathbf{sh}_1^p - c_1^{op} \tau^{op} \Delta \mathbf{sh}_1^{op}$$

et du bénéfice futur

$$\mathbb{E}_{\mathbf{A}_2|A_1} \left[ \max_{\substack{\mathbf{X}_2, \mathbf{V}_2 \\ \mathbf{V}_2=f(\mathbf{V}_1, \mathbf{A}_2, \mathbf{X}_2)}} B_2(\mathbf{V}_1, \mathbf{A}_2, \mathbf{X}_2) + \dots \right].$$

Pour une semaine  $t = 2 \dots T-1$ , partant d'un état initial du système  $\mathbf{V}_{t-1}$  et sachant qu'il y aura un apport hydrique  $\mathbf{A}_t$ , l'objectif est de prendre une décision  $\mathbf{X}_t$  de sorte à maximiser la somme du bénéfice immédiat

$$B_t(\mathbf{V}_{t-1}, \mathbf{A}_t, \mathbf{X}_t) = \nu_t^p \tau^p \Delta \mathbf{p}_t^p + \nu_t^{op} \tau^{op} \Delta \mathbf{p}_t^{op} - c_t^p \tau^p \Delta \mathbf{sh}_t^p - c_t^{op} \tau^{op} \Delta \mathbf{sh}_t^{op}$$

et du bénéfice futur

$$\mathbb{E}_{\mathbf{A}_{t+1}|\mathbf{A}_t} \left[ \max_{\substack{\mathbf{X}_{t+1}, \mathbf{V}_{t+1} \\ \mathbf{V}_{t+1}=f(\mathbf{V}_t, \mathbf{A}_{t+1}, \mathbf{X}_{t+1})}} B_{t+1}(\mathbf{V}_t, \mathbf{A}_{t+1}, \mathbf{X}_{t+1}) + \dots \right].$$

Pour  $t = T$ , partant d'un état initial du système  $\mathbf{V}_{T-1}$  et sachant qu'il y aura un apport hydrique  $\mathbf{A}_T$ , l'objectif est de prendre une décision  $\mathbf{X}_T$  de sorte à maximiser le bénéfice immédiat

$$B_T(\mathbf{V}_{T-1}, \mathbf{A}_T, \mathbf{X}_T) = \nu_T^p \tau^p \Delta \mathbf{p}_T^p + \nu_T^{op} \tau^{op} \Delta \mathbf{p}_T^{op} - c_T^p \tau^p \Delta \mathbf{sh}_T^p - c_T^{op} \tau^{op} \Delta \mathbf{sh}_T^{op}.$$

Le bénéfice immédiat correspond au profit tiré de la production de la semaine considérée. Par exemple, le terme  $\nu_t^p \tau^p \Delta \mathbf{p}_t^p$  est le prix de l'électricité fournie durant la demande de pointe. En effet,  $\mathbf{p}_t^p$  étant la puissance instantanée (MW) fournie,  $\tau^p \Delta$  étant le nombre d'heures de la classe de demande de pointe, on obtient la quantité fournie  $\tau^p \Delta \mathbf{p}_t^p$  de MWh que l'on multiplie par le prix  $\nu_t^p$  du MWh. Les autres termes de l'objectif peuvent être interprétés de manière similaire.



### 4.3.2 Contraintes

Les contraintes  $\mathbf{V}_t = f(\mathbf{V}_{t-1}, \mathbf{A}_t, \mathbf{X}_t)$  relient le volume finale  $\mathbf{V}_t$  au volume initial  $\mathbf{V}_{t-1}$ , à l'apport aléatoire  $\mathbf{A}_t$  et aux décisions  $\mathbf{X}_t$ .

#### Contraintes liées aux réservoirs

Les contraintes de la forme

$$\begin{aligned} \mathbf{v}_{r,t} = \mathbf{v}_{r,t-1} &- \sum_{c \in \mathcal{C}^d(r)} \left( \tau^p \rho \mathbf{q}_{c,t}^p + \tau^{op} \rho \mathbf{q}_{c,t}^{op} \right) - \sum_{oe \in \mathcal{OE}^d(r)} \rho \mathbf{q}_{oe,t} \\ &+ \sum_{c \in \mathcal{C}^u(r)} \left( \tau^p \rho \mathbf{q}_{c,t}^p + \tau^{op} \rho \mathbf{q}_{c,t}^{op} \right) + \sum_{oe \in \mathcal{OE}^u(r)} \rho \mathbf{q}_{oe,t} + \delta_{r,t} \rho \mathbf{A}_t \\ \sum_{c \in \mathcal{C}^d(r)} \left( \tau^p \rho \mathbf{q}_{c,t}^p + \tau^{op} \rho \mathbf{q}_{c,t}^{op} \right) &+ \sum_{oe \in \mathcal{OE}^d(r)} \rho \mathbf{q}_{oe,t} = \sum_{c \in \mathcal{C}^u(r)} \left( \tau^p \rho \mathbf{q}_{c,t}^p + \tau^{op} \rho \mathbf{q}_{c,t}^{op} \right) + \sum_{oe \in \mathcal{OE}^u(r)} \rho \mathbf{q}_{oe,t} + \delta_{r,t} \rho \mathbf{A}_t \end{aligned}$$

correspondent aux équations de conservation de masse des réservoirs à capacité positive et des réservoirs sans capacité, respectivement.

Pour un réservoir à capacité positive  $r$ , la contrainte

$$\mathbf{l}_{r,t}^u = l_r \left( \frac{\mathbf{v}_{r,t-1} + \mathbf{v}_{r,t}}{2} \right)$$

donne le niveau du réservoir  $r$  en fonction de son volume moyen (sur une semaine). La fonction  $l_r(\cdot)$  est concave linéaire par morceaux, La manière de la représenter dans un modèle linéaire est détaillée dans l'annexe A.

La contrainte

$$0 \leq \mathbf{v}_{r,t} \leq \overline{v}_r$$

permet de respecter la capacité de stockage maximale du réservoir à capacité positive  $r$ .

Pour un ouvrage d'évacuation  $oe$  lié à un réservoir  $r$ , la contrainte

$$0 \leq \mathbf{q}_{oe,t} \leq \overline{q}_{oe}$$

permet de respecter sa capacité d'évacuation.

### Contraintes liées aux centrales

Pour une centrale  $c$  à hauteur de chute variable, la contrainte

$$\mathbf{h}_{c,t} = \mathbf{l}_{r(c),t}^u - l_c^d$$

donne la hauteur de chute de la centrale  $c$ , qui est la différence entre le niveau en amont  $\mathbf{l}_{r(c),t}^u$  (niveau du réservoir  $r(c)$ ) et le niveau en aval  $l_c^d$ .

Les contraintes

$$\begin{cases} \mathbf{p}_{c,t}^p \leq a_{c,q}^{(j)} \mathbf{q}_{c,t}^p + a_{c,h}^{(j)} \mathbf{h}_{c,t} + b_c^{(j)}, & j = 1 \dots K_c \\ \mathbf{p}_{c,t}^{op} \leq a_{c,q}^{(j)} \mathbf{q}_{c,t}^{op} + a_{c,h}^{(j)} \mathbf{h}_{c,t} + b_c^{(j)}, & j = 1 \dots K_c \end{cases}$$

donnent la puissance générée par la centrale à hauteur de chute variable  $c$  durant la classe de demande de pointe et la classe de demande moyenne, respectivement. Il s'agit d'une approximation concave linéaire par morceaux de la fonction de production de la centrale  $c$  (voir section 3.2.4).

Les contraintes

$$0 \leq \mathbf{q}_{c,t}^p, \mathbf{q}_{c,t}^{op} \leq \frac{\mathbf{h}_{c,t}}{h_c^0} \overline{q_c^0}$$

donnent les bornes sur le débit turbiné de la centrale : la borne supérieure de débit dépend de la hauteur de chute.

Pour une centrale  $c$  à hauteur de chute fixe, les contraintes

$$\begin{cases} \mathbf{p}_{c,t}^p = p_c(\mathbf{q}_{c,t}^p) \\ \mathbf{p}_{c,t}^{op} = p_c(\mathbf{q}_{c,t}^{op}) \end{cases}$$

donnent la puissance générée de la centrale durant la classe de demande de pointe et la classe de demande moyenne, respectivement. La fonction  $p_c(\cdot)$  est concave linéaire par morceaux, La manière de la représenter dans un modèle linéaire est détaillée dans l'annexe A. La contrainte

$$0 \leq \mathbf{q}_{c,t}^p, \mathbf{q}_{c,t}^{op} \leq \overline{q_c^0}$$

donne les bornes sur le débit turbiné de la centrale  $c$ .

### Contraintes d'équilibre électrique

Les contraintes

$$\begin{cases} \mathbf{p}_t^p = \sum_{c \in \mathcal{C}} \mathbf{p}_{c,t}^p \\ \mathbf{p}_t^{op} = \sum_{c \in \mathcal{C}} \mathbf{p}_{c,t}^{op} \end{cases}$$

expriment la puissance instantanée totale générée pour la classe de demande de pointe et la classe de demande moyenne, respectivement.

Finalement, les contraintes

$$\begin{cases} \mathbf{p}_t^p + \mathbf{sh}_t^p \geq d_t^p \\ \mathbf{p}_t^{op} + \mathbf{sh}_t^{op} \geq d_t^{op} \end{cases}$$

correspondent à l'équilibre offre-demande pour la classe de demande de pointe et la classe de demande moyenne, respectivement. Il est à noter que l'on se permet de produire plus que la demande si l'on peut.

Ci-après une formulation plus détaillée du modèle avec les variables, contraintes et objectifs.

$$\max_{\substack{\mathbf{X}_1, \mathbf{V}_1 \\ \mathbf{V}_1=f(\mathbf{V}_0, \mathbf{A}_1, \mathbf{X}_1)}} B_1(V_0, A_1, \mathbf{X}_1) + \mathbb{E}_{\mathbf{A}_2|A_1} \left[ \max_{\substack{\mathbf{X}_2, \mathbf{V}_2 \\ \mathbf{V}_2=f(\mathbf{V}_1, \mathbf{A}_2, \mathbf{X}_2)}} B_2(\mathbf{V}_1, \mathbf{A}_2, \mathbf{X}_2) + \dots \right] \text{équivaut à}$$

$$\max_{\mathbf{X}_1, \mathbf{V}_1} \nu_1^p \tau^p \Delta \mathbf{p}_1^p + \nu_1^{op} \tau^{op} \Delta \mathbf{p}_1^{op} - c_1^p \tau^p \Delta \mathbf{sh}_1^p - c_1^{op} \tau^{op} \Delta \mathbf{sh}_1^{op} +$$

$$\mathbb{E}_{\mathbf{A}_2|A_1} \left[ \max_{\substack{\mathbf{X}_2, \mathbf{V}_2 \\ \mathbf{V}_2=f(\mathbf{V}_1, \mathbf{A}_2, \mathbf{X}_2)}} B_2(\mathbf{V}_1, \mathbf{A}_2, \mathbf{X}_2) + \dots \right]$$

$$\begin{aligned} \text{s. à } \mathbf{v}_{r,1} &= v_{r,0} - \sum_{c \in \mathcal{C}^d(r)} \left( \tau^p \rho \mathbf{q}_{c,1}^p + \tau^{op} \rho \mathbf{q}_{c,1}^{op} \right) - \sum_{oe \in \mathcal{OE}^d(r)} \rho \mathbf{q}_{oe,1} \\ &+ \sum_{c \in \mathcal{C}^u(r)} \left( \tau^p \rho \mathbf{q}_{c,1}^p + \tau^{op} \rho \mathbf{q}_{c,1}^{op} \right) + \sum_{oe \in \mathcal{OE}^u(r)} \rho \mathbf{q}_{oe,1} + \delta_{r,1} \rho A_1 \quad \forall r \in \mathcal{R}^+ \end{aligned}$$

$$\begin{aligned} \sum_{c \in \mathcal{C}^d(r)} \left( \tau^p \rho \mathbf{q}_{c,1}^p + \tau^{op} \rho \mathbf{q}_{c,1}^{op} \right) + \sum_{oe \in \mathcal{OE}^d(r)} \rho \mathbf{q}_{oe,1} &= \\ \sum_{c \in \mathcal{C}^u(r)} \left( \tau^p \rho \mathbf{q}_{c,1}^p + \tau^{op} \rho \mathbf{q}_{c,1}^{op} \right) + \sum_{oe \in \mathcal{OE}^u(r)} \rho \mathbf{q}_{oe,1} + \delta_{r,1} \rho A_1 &\quad \forall r \in \mathcal{R}^- \end{aligned}$$

$$\mathbf{l}_{r,1}^u = l_r \left( \frac{v_{r,0} + \mathbf{v}_{r,1}}{2} \right) \quad \forall r \in \mathcal{R}^+$$

$$\mathbf{h}_{c,1} = \mathbf{l}_{r(c),1}^u - l_c^d \quad \forall c \in \mathcal{C}^+$$

$$\mathbf{p}_{c,1}^p \leq a_{c,q}^{(j)} \mathbf{q}_{c,1}^p + a_{c,h}^{(j)} \mathbf{h}_{c,1} + b_c^{(j)} \quad \forall c \in \mathcal{C}^+, j = 1 \dots K_c$$

$$\mathbf{p}_{c,1}^{op} \leq a_{c,q}^{(j)} \mathbf{q}_{c,1}^{op} + a_{c,h}^{(j)} \mathbf{h}_{c,1} + b_c^{(j)} \quad \forall c \in \mathcal{C}^+, j = 1 \dots K_c$$

$$\mathbf{p}_{c,1}^p = p_c \left( \mathbf{q}_{c,1}^p \right) \quad \forall c \in \mathcal{C}^-$$

$$\mathbf{p}_{c,1}^{op} = p_c \left( \mathbf{q}_{c,1}^{op} \right) \quad \forall c \in \mathcal{C}^-$$

$$\mathbf{p}_1^p = \sum_{c \in \mathcal{C}} \mathbf{p}_{c,1}^p$$

$$\mathbf{p}_1^{op} = \sum_{c \in \mathcal{C}} \mathbf{p}_{c,1}^{op}$$

$$\mathbf{p}_1^p + \mathbf{sh}_1^p \geq d_1^p$$

$$\mathbf{p}_1^{op} + \mathbf{sh}_1^{op} \geq d_1^{op}$$

$$0 \leq \mathbf{v}_{r,1} \leq \overline{v_r} \quad \forall r \in \mathcal{R}^+$$

$$0 \leq \mathbf{q}_{c,1}^p, \mathbf{q}_{c,1}^{op} \leq \frac{\mathbf{h}_{c,1}}{h_c^0} \overline{q_c^0} \quad \forall c \in \mathcal{C}^+$$

$$0 \leq \mathbf{q}_{c,1}^p, \mathbf{q}_{c,1}^{op} \leq \overline{q_c^0} \quad \forall c \in \mathcal{C}^-$$

$$0 \leq \mathbf{q}_{oe,1} \leq \overline{q_{oe}} \quad \forall oe \in \mathcal{OE}$$

Pour  $t = 2 \dots T - 1$

$$\max_{\substack{\mathbf{X}_t, \mathbf{V}_t \\ \mathbf{V}_t = f(\mathbf{V}_{t-1}, \mathbf{A}_t, \mathbf{X}_t)}} B_t(\mathbf{V}_{t-1}, \mathbf{A}_t, \mathbf{X}_t) + \mathbb{E}_{\mathbf{A}_{t+1}|\mathbf{A}_t} \left[ \max_{\substack{\mathbf{X}_{t+1}, \mathbf{V}_{t+1} \\ \mathbf{V}_{t+1} = f(\mathbf{V}_t, \mathbf{A}_{t+1}, \mathbf{X}_{t+1})}} B_{t+1}(\mathbf{V}_t, \mathbf{A}_{t+1}, \mathbf{X}_{t+1}) + \dots \right]$$

équivalent à

$$\begin{aligned} & \max_{\mathbf{X}_t, \mathbf{V}_t} \nu_t^p \tau^p \Delta \mathbf{p}_t^p + \nu_t^{op} \tau^{op} \Delta \mathbf{p}_t^{op} - c_t^p \tau^p \Delta \mathbf{sh}_t^p - c_t^{op} \tau^{op} \Delta \mathbf{sh}_t^{op} + \\ & \quad \mathbb{E}_{\mathbf{A}_{t+1}|\mathbf{A}_t} \left[ \max_{\substack{\mathbf{X}_{t+1}, \mathbf{V}_{t+1} \\ \mathbf{V}_{t+1} = f(\mathbf{V}_t, \mathbf{A}_{t+1}, \mathbf{X}_{t+1})}} B_{t+1}(\mathbf{V}_t, \mathbf{A}_{t+1}, \mathbf{X}_{t+1}) + \dots \right] \\ \text{s. à } & \mathbf{v}_{r,t} = \mathbf{v}_{r,t-1} - \sum_{c \in \mathcal{C}^d(r)} \left( \tau^p \rho \mathbf{q}_{c,t}^p + \tau^{op} \rho \mathbf{q}_{c,t}^{op} \right) - \sum_{oe \in \mathcal{OE}^d(r)} \rho \mathbf{q}_{oe,t} \\ & \quad + \sum_{c \in \mathcal{C}^u(r)} \left( \tau^p \rho \mathbf{q}_{c,t}^p + \tau^{op} \rho \mathbf{q}_{c,t}^{op} \right) + \sum_{oe \in \mathcal{OE}^u(r)} \rho \mathbf{q}_{oe,t} + \delta_{r,t} \rho \mathbf{A}_t \quad \forall r \in \mathcal{R}^+ \\ & \quad \sum_{c \in \mathcal{C}^d(r)} \left( \tau^p \rho \mathbf{q}_{c,t}^p + \tau^{op} \rho \mathbf{q}_{c,t}^{op} \right) + \sum_{oe \in \mathcal{OE}^d(r)} \rho \mathbf{q}_{oe,t} = \\ & \quad \sum_{c \in \mathcal{C}^u(r)} \left( \tau^p \rho \mathbf{q}_{c,t}^p + \tau^{op} \rho \mathbf{q}_{c,t}^{op} \right) + \sum_{oe \in \mathcal{OE}^u(r)} \rho \mathbf{q}_{oe,t} + \delta_{r,t} \rho \mathbf{A}_t \quad \forall r \in \mathcal{R}^- \\ & \mathbf{l}_{r,t}^u = l_r \left( \frac{\mathbf{v}_{r,t-1} + \mathbf{v}_{r,t}}{2} \right) \quad \forall r \in \mathcal{R}^+ \\ & \mathbf{h}_{c,t} = \mathbf{l}_{r(c),t}^u - l_c^d \quad \forall c \in \mathcal{C}^+ \\ & \mathbf{p}_{c,t}^p \leq a_{c,q}^{(j)} \mathbf{q}_{c,t}^p + a_{c,h}^{(j)} \mathbf{h}_{c,t} + b_c^{(j)} \quad \forall c \in \mathcal{C}^+, j = 1 \dots K_c \\ & \mathbf{p}_{c,t}^{op} \leq a_{c,q}^{(j)} \mathbf{q}_{c,t}^{op} + a_{c,h}^{(j)} \mathbf{h}_{c,t} + b_c^{(j)} \quad \forall c \in \mathcal{C}^+, j = 1 \dots K_c \\ & \mathbf{p}_{c,t}^p = p_c \left( \mathbf{q}_{c,t}^p \right) \quad \forall c \in \mathcal{C}^- \\ & \mathbf{p}_{c,t}^{op} = p_c \left( \mathbf{q}_{c,t}^{op} \right) \quad \forall c \in \mathcal{C}^- \\ & \mathbf{p}_t^p = \sum_{c \in \mathcal{C}} \mathbf{p}_{c,t}^p \\ & \mathbf{p}_t^{op} = \sum_{c \in \mathcal{C}} \mathbf{p}_{c,t}^{op} \\ & \mathbf{p}_t^p + \mathbf{sh}_t^p \geq d_t^p \\ & \mathbf{p}_t^{op} + \mathbf{sh}_t^{op} \geq d_t^{op} \\ & 0 \leq \mathbf{v}_{r,t} \leq \overline{v_r} \quad \forall r \in \mathcal{R}^+ \\ & 0 \leq \mathbf{q}_{c,t}^p, \mathbf{q}_{c,t}^{op} \leq \frac{\mathbf{h}_{c,t}}{h_c^0} \overline{q_c^0} \quad \forall c \in \mathcal{C}^+ \\ & 0 \leq \mathbf{q}_{c,t}^p, \mathbf{q}_{c,t}^{op} \leq \overline{q_c^0} \quad \forall c \in \mathcal{C}^- \\ & 0 \leq \mathbf{q}_{oe,t} \leq \overline{q_{oe}} \quad \forall oe \in \mathcal{OE} \end{aligned}$$

$$\max_{\substack{\mathbf{X}_T, \mathbf{V}_T \\ \mathbf{V}_T = f(\mathbf{V}_{T-1}, \mathbf{A}_T, \mathbf{X}_T)}} B_T(\mathbf{V}_{T-1}, \mathbf{A}_T, \mathbf{X}_T) \quad \text{équivaut à}$$

$$\max_{\mathbf{X}_T, \mathbf{V}_T} \nu_T^p \tau^p \Delta \mathbf{p}_T^p + \nu_T^{op} \tau^{op} \Delta \mathbf{p}_T^{op} - c_T^p \tau^p \Delta \mathbf{sh}_T^p - c_T^{op} \tau^{op} \Delta \mathbf{sh}_T^{op}$$

$$\text{s. à } \mathbf{v}_{r,T} = \mathbf{v}_{r,T-1} - \sum_{c \in \mathcal{C}^d(r)} \left( \tau^p \rho \mathbf{q}_{c,T}^p + \tau^{op} \rho \mathbf{q}_{c,T}^{op} \right) - \sum_{oe \in \mathcal{OE}^d(r)} \rho \mathbf{q}_{oe,T}$$

$$+ \sum_{c \in \mathcal{C}^u(r)} \left( \tau^p \rho \mathbf{q}_{c,T}^p + \tau^{op} \rho \mathbf{q}_{c,T}^{op} \right) + \sum_{oe \in \mathcal{OE}^u(r)} \rho \mathbf{q}_{oe,T} + \delta_{r,T} \rho \mathbf{A}_T \quad \forall r \in \mathcal{R}^+$$

$$\sum_{c \in \mathcal{C}^d(r)} \left( \tau^p \rho \mathbf{q}_{c,T}^p + \tau^{op} \rho \mathbf{q}_{c,T}^{op} \right) + \sum_{oe \in \mathcal{OE}^d(r)} \rho \mathbf{q}_{oe,T} =$$

$$\sum_{c \in \mathcal{C}^u(r)} \left( \tau^p \rho \mathbf{q}_{c,T}^p + \tau^{op} \rho \mathbf{q}_{c,T}^{op} \right) + \sum_{oe \in \mathcal{OE}^u(r)} \rho \mathbf{q}_{oe,T} + \delta_{r,T} \rho \mathbf{A}_T \quad \forall r \in \mathcal{R}^-$$

$$\mathbf{l}_{r,T}^u = l_r \left( \frac{\mathbf{v}_{r,T-1} + \mathbf{v}_{r,T}}{2} \right) \quad \forall r \in \mathcal{R}^+$$

$$\mathbf{h}_{c,T} = \mathbf{l}_{r(c),T}^u - l_c^d \quad \forall c \in \mathcal{C}^+$$

$$\mathbf{p}_{c,T}^p \leq a_{c,q}^{(j)} \mathbf{q}_{c,T}^p + a_{c,h}^{(j)} \mathbf{h}_{c,T} + b_c^{(j)} \quad \forall c \in \mathcal{C}^+, j = 1 \dots K_c$$

$$\mathbf{p}_{c,T}^{op} \leq a_{c,q}^{(j)} \mathbf{q}_{c,T}^{op} + a_{c,h}^{(j)} \mathbf{h}_{c,T} + b_c^{(j)} \quad \forall c \in \mathcal{C}^+, j = 1 \dots K_c$$

$$\mathbf{p}_{c,T}^p = p_c \left( \mathbf{q}_{c,T}^p \right) \quad \forall c \in \mathcal{C}^-$$

$$\mathbf{p}_{c,T}^{op} = p_c \left( \mathbf{q}_{c,T}^{op} \right) \quad \forall c \in \mathcal{C}^-$$

$$\mathbf{p}_T^p = \sum_{c \in \mathcal{C}} \mathbf{p}_{c,T}^p$$

$$\mathbf{p}_T^{op} = \sum_{c \in \mathcal{C}} \mathbf{p}_{c,T}^{op}$$

$$\mathbf{p}_T^p + \mathbf{sh}_T^p \geq d_T^p$$

$$\mathbf{p}_T^{op} + \mathbf{sh}_T^{op} \geq d_T^{op}$$

$$0 \leq \mathbf{v}_{r,T} \leq \overline{v_r} \quad \forall r \in \mathcal{R}^+$$

$$0 \leq \mathbf{q}_{c,T}^p, \mathbf{q}_{c,T}^{op} \leq \frac{\mathbf{h}_{c,T}}{h_c^0} \overline{q_c^0} \quad \forall c \in \mathcal{C}^+$$

$$0 \leq \mathbf{q}_{c,T}^p, \mathbf{q}_{c,T}^{op} \leq \overline{q_c^0} \quad \forall c \in \mathcal{C}^-$$

$$0 \leq \mathbf{q}_{oe,T} \leq \overline{q_{oe}} \quad \forall oe \in \mathcal{OE}$$

## CHAPITRE 5 MÉTHODES DE RÉOLUTION PAR PROGRAMMATION DYNAMIQUE STOCHASTIQUE

Ce chapitre présente des résultats théoriques et la méthodologie utilisée pour résoudre le problème de planification moyen-terme de la production d'hydroélectricité. En particulier, on reformule le problème de planification (4.1) en un schéma de programmation dynamique. Ensuite, on expose une méthode exacte de programmation dynamique (SDP) pour le résoudre dans sa nouvelle forme et on montre que cette méthode a une complexité exponentielle. Deux méthodes qui dérivent de SDP : SDDP et ASDP, et qui essaient de contourner sa complexité sont présentées par la suite.

Il est à noter que la programmation dynamique présentée et utilisée dans ce mémoire est la programmation dynamique **duale**. Par conséquent, on majore les fonctions de valeur de Bellman par leurs hyperplans tangents ou coupants. Ainsi, serait-il plus cohérent d'appeler les méthodes présentées dans ce chapitre respectivement par *Programmation dynamique duale stochastique* (SDDP), *Programmation dynamique duale stochastique itérative* (ISDDP) et *Programmation dynamique duale stochastique approchée* (ASDDP). Cependant, le fait que la deuxième méthode est largement connue sous le nom SDDP nous a incité à adopter les noms respectifs de SDP, SDDP et ASDP.

### 5.1 Notation

En plus de la notation présentée dans le chapitre précédent, on introduit la notation supplémentaire suivante

#### Processus stochastique d'apports

Symbole	Unité	Description
$\tilde{\mathbf{A}}_t$	m <sup>3</sup> /s	variable aléatoire (discrète) de l'apport à la rivière durant la semaine $t$
$N$		nombre de valeurs discrètes des variables aléatoires $\tilde{\mathbf{A}}$
$\tilde{A}_t^{(i)}$	m <sup>3</sup> /s	$i^{\text{e}}$ valeur discrète de la variable aléatoire $\tilde{\mathbf{A}}_t$
$P_{ij}^t$		probabilité de transition de $\tilde{A}_{t-1}^{(i)}$ , durant la semaine $t - 1$ , à $\tilde{A}_t^{(j)}$ , durant la semaine $t$
$i_1$		indice de $A_1$ dans la discrétisation de la variable aléatoire d'apport $\tilde{\mathbf{A}}_1$

## Symboles

Symbole	Description
$\mathcal{F}_t$	fonction de bénéfice futur de la semaine $t-1$ (cas du vrai processus stochastique d'apport $(\mathbf{A}_t)_t$ )
$\tilde{\mathcal{F}}_t$	fonction de bénéfice futur de la semaine $t-1$ (cas du processus stochastique d'apport discrétisé $(\tilde{\mathbf{A}}_t)_t$ )
$\tilde{\mathcal{F}}_t^{(i)}$	fonction de bénéfice futur de la semaine $t-1$ restreinte à l'apport $\tilde{A}_{t-1}^{(i)}$
$\tilde{\mathfrak{F}}_t^{(i)}$	approximation de la fonction $\tilde{\mathcal{F}}_t^{(i)}$

## 5.2 Reformulation du problème : schéma de programmation dynamique

Définissons les formules de récurrence suivantes

$$\begin{aligned}
 \Theta_1 &= \max_{\substack{\mathbf{X}_1, \mathbf{V}_1 \\ \mathbf{V}_1 = f(V_0, A_1, \mathbf{X}_1)}} B_1(V_0, A_1, \mathbf{X}_1) + \mathcal{F}_2(A_1, \mathbf{V}_1) \\
 \mathcal{F}_t(A_{t-1}, V_{t-1}) &= \mathbb{E}_{\mathbf{A}_t | A_{t-1}} \left[ \max_{\substack{\mathbf{X}_t, \mathbf{V}_t \\ \mathbf{V}_t = f(V_{t-1}, A_t, \mathbf{X}_t)}} B_t(V_{t-1}, A_t, \mathbf{X}_t) + \mathcal{F}_{t+1}(\mathbf{A}_t, \mathbf{V}_t) \right] \quad t = 2 \dots T \\
 \mathcal{F}_{T+1}(A_T, V_T) &\equiv 0.
 \end{aligned} \tag{5.1}$$

Les équations (5.1) constituent le schéma de programmation dynamique équivalent au problème (4.1).

la fonction  $\mathcal{F}_{t+1}$  est appelée fonction de bénéfice futur de la semaine  $t$ . Elle valorise l'état du système à la fin de la semaine  $t$ . Ici, puisque les apports hydriques sont temporellement corrélés (modèle PAR(1)), l'état du système est constitué, en plus des volumes des réservoirs, de l'apport passé.

L'objectif de la planification moyen-terme de la production d'hydroélectricité est de déterminer ces fonctions de bénéfice futur qui permettront de prendre des décisions optimales à chaque semaine qui tiennent compte de l'aléa futur. Or, il est difficile de trouver ces fonctions lorsque les distributions des variables aléatoires  $(\mathbf{A}_t)_t$  sont continues (nombre infini de fonctions à déterminer). C'est pour cette raison que l'on discrétise le processus d'apport  $(\mathbf{A}_t)_t$  comme développé dans la section 3.3.3 et l'on passe du processus d'apport continue  $(\mathbf{A}_t)_t$  au



processus d'apport discret  $(\tilde{\mathbf{A}}_t)_t$ . Les formules de récurrence précédentes deviennent

$$\begin{aligned}\tilde{\Theta}_1 &= \max_{\substack{\mathbf{X}_1, \mathbf{V}_1 \\ \mathbf{V}_1 = f(V_0, A_1, \mathbf{X}_1)}} B_1(V_0, A_1, \mathbf{X}_1) + \tilde{\mathcal{F}}_2(A_1, \mathbf{V}_1) \\ \tilde{\mathcal{F}}_t(\tilde{A}_{t-1}, V_{t-1}) &= \mathbb{E}_{\tilde{\mathbf{A}}_t | \tilde{A}_{t-1}} \left[ \max_{\substack{\mathbf{X}_t, \mathbf{V}_t \\ \mathbf{V}_t = f(V_{t-1}, \tilde{\mathbf{A}}_t, \mathbf{X}_t)}} B_t(V_{t-1}, \tilde{\mathbf{A}}_t, \mathbf{X}_t) + \tilde{\mathcal{F}}_{t+1}(\tilde{\mathbf{A}}_t, \mathbf{V}_t) \right] \quad t = 2 \dots T \\ \tilde{\mathcal{F}}_{T+1} &\equiv 0.\end{aligned}\tag{5.2}$$

### 5.3 Programmation dynamique stochastique et politique de gestion

Si les décisions à chaque semaine  $t$  sont prises en utilisant les fonctions de bénéfice futur  $(\mathcal{F}_t)_t$ , on dit que les décisions suivent une politique de gestion. En particuliers, le vecteur des volumes des réservoirs à la fin de la semaine  $t$ ,  $V_t$ , est une fonction des apports qui ont eu lieu jusqu'à  $t$ , i.e.  $V_t = V_t(A_1, \mathbf{A}_2, \dots, \mathbf{A}_t)$ . La figure suivante illustre cet aspect

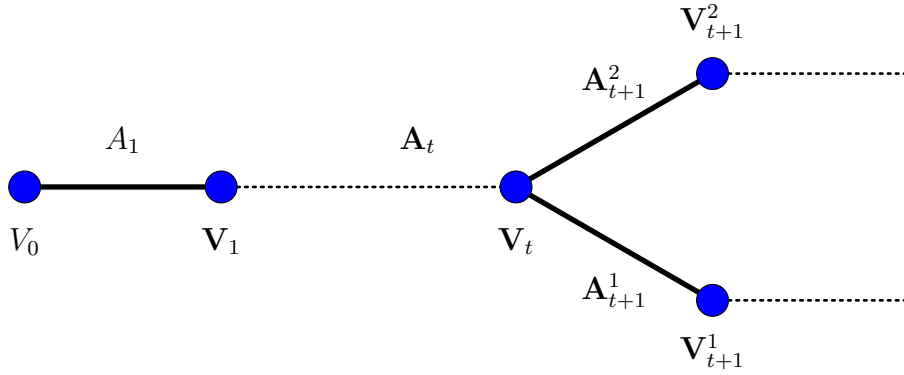


Figure 5.1 Deux scénarios partageant les mêmes apports jusqu'à la semaine  $t$

Les deux scénarios de la figure partageront les mêmes décisions, i.e. les mêmes volumes de réservoirs  $\mathbf{V}_1, \dots, \mathbf{V}_t$ .

Cette propriété de politique de gestion nous permet d'écrire

$$\Theta_1 = \mathbb{E}_{\mathbf{A}_2, \dots, \mathbf{A}_T | A_1} \left[ B_1(V_0, A_1, \mathbf{X}_1) + B_2(\mathbf{V}_1, \mathbf{A}_2, \mathbf{X}_2) + \dots B_T(\mathbf{V}_{T-1}, \mathbf{A}_T, \mathbf{X}_T) \right] \tag{5.3}$$

et

$$\tilde{\Theta}_1 = \mathbb{E}_{\tilde{\mathbf{A}}_2, \dots, \tilde{\mathbf{A}}_T | A_1} \left[ B_1(V_0, A_1, \mathbf{X}_1) + B_2(\mathbf{V}_1, \tilde{\mathbf{A}}_2, \mathbf{X}_2) + \dots + B_T(\mathbf{V}_{T-1}, \tilde{\mathbf{A}}_T, \mathbf{X}_T) \right]. \quad (5.4)$$

#### 5.4 Concavité des fonctions de bénéfice futur

Dans cette section, on démontre que la fonction de bénéfice futur  $\tilde{\mathcal{F}}_t$  est concave et linéaire par morceaux par rapport à  $V_{t-1}$ . On démontre aussi que cette propriété n'est pas vérifiée pour  $A_{t-1}$ .

##### 5.4.1 Concavité par rapport à l'état *volume*

Dans Birge et Louveaux (2011), il est prouvé, dans le cas de variables aléatoires continues indépendantes, que les fonctions de bénéfice futur sont concaves par rapport à la variable d'état et que, lorsque le support des variables aléatoires est fini, ces fonctions sont de plus linéaires par morceaux. La preuve est généralisable au cas que l'on étudie. En effet, pour  $t = T$ ,

$$\tilde{\mathcal{F}}_T(\tilde{A}_{T-1}, V_{T-1}) = \mathbb{E}_{\tilde{\mathbf{A}}_T | \tilde{A}_{T-1}} \left[ \max_{\substack{\mathbf{X}_T, \mathbf{V}_T \\ \mathbf{V}_T = f(V_{T-1}, \tilde{\mathbf{A}}_T, \mathbf{X}_T)}} B_T(V_{T-1}, \tilde{\mathbf{A}}_T, \mathbf{X}_T) \right] \quad (\tilde{\mathcal{F}}_{T+1} \equiv 0)$$

Or  $(\tilde{A}_{T-1}, V_{T-1}) \mapsto \max_{\substack{\mathbf{X}_T, \mathbf{V}_T \\ \mathbf{V}_T = f(V_{T-1}, \tilde{\mathbf{A}}_T, \mathbf{X}_T)}} B_T(V_{T-1}, \tilde{\mathbf{A}}_T, \mathbf{X}_T)$  est concave linéaire par morceaux par rapport à  $V_{T-1}$  car  $V_{T-1}$  fait partie du second membre des contraintes  $\mathbf{V}_T = f(V_{T-1}, \tilde{\mathbf{A}}_T, \mathbf{X}_T)$  et ainsi  $(\tilde{A}_{T-1}, V_{T-1}) \mapsto \tilde{\mathcal{F}}_T(\tilde{A}_{T-1}, V_{T-1})$  est aussi concave linéaire par morceaux par rapport à  $V_{T-1}$  car l'espérance mathématique est sur un ensemble fini de valeurs possibles de la variable aléatoire  $\tilde{\mathbf{A}}_T$ .

Pour  $t = T - 1$ ,

$$\tilde{\mathcal{F}}_{T-1}(\tilde{A}_{T-2}, V_{T-2}) = \mathbb{E}_{\tilde{\mathbf{A}}_{T-1} | \tilde{A}_{T-2}} \left[ \max_{\substack{\mathbf{X}_{T-1}, \mathbf{V}_{T-1} \\ \mathbf{V}_{T-1} = f(V_{T-2}, \tilde{\mathbf{A}}_{T-1}, \mathbf{X}_{T-1})}} B_{T-1}(V_{T-2}, \tilde{\mathbf{A}}_{T-1}, \mathbf{X}_{T-1}) + \tilde{\mathcal{F}}_T(\tilde{\mathbf{A}}_{T-1}, \mathbf{V}_{T-1}) \right].$$

Puisque  $(\tilde{\mathbf{A}}_{T-1}, V_{T-1}) \mapsto \tilde{\mathcal{F}}_T(\tilde{\mathbf{A}}_{T-1}, V_{T-1})$  est concave linéaire par morceaux par rapport à

$V_{T-1}$ , le modèle

$$\max_{\substack{\mathbf{X}_{T-1}, \mathbf{V}_{T-1} \\ \mathbf{V}_{T-1}=f(V_{T-2}, \tilde{\mathbf{A}}_{T-1}, \mathbf{X}_{T-1})}} B_{T-1}(V_{T-2}, \tilde{\mathbf{A}}_{T-1}, \mathbf{X}_{T-1}) + \tilde{\mathcal{F}}_T(\tilde{\mathbf{A}}_{T-1}, \mathbf{V}_{T-1})$$

est linéaire, alors

$$(\tilde{A}_{T-2}, V_{T-2}) \mapsto \max_{\substack{\mathbf{X}_{T-1}, \mathbf{V}_{T-1} \\ \mathbf{V}_{T-1}=f(V_{T-2}, \tilde{\mathbf{A}}_{T-1}, \mathbf{X}_{T-1})}} B_{T-1}(V_{T-2}, \tilde{\mathbf{A}}_{T-1}, \mathbf{X}_{T-1}) + \tilde{\mathcal{F}}_T(\tilde{\mathbf{A}}_{T-1}, \mathbf{V}_{T-1})$$

est concave linéaire par morceaux par rapport à  $V_{T-2}$  car  $V_{T-2}$  fait partie du second membre des contraintes  $\mathbf{V}_{T-1} = f(V_{T-2}, \tilde{\mathbf{A}}_{T-1}, \mathbf{X}_{T-1})$  et ainsi  $(\tilde{A}_{T-2}, V_{T-2}) \mapsto \tilde{\mathcal{F}}_{T-1}(\tilde{A}_{T-2}, V_{T-2})$  est aussi concave linéaire par morceaux par rapport à  $V_{T-2}$  car l'espérance mathématique est sur un ensemble fini de valeurs possibles de la variable aléatoire  $\tilde{\mathbf{A}}_{T-1}$ .

Il suffit de répéter le même raisonnement pour montrer que  $(\tilde{A}_{t-1}, V_{t-1}) \mapsto \tilde{\mathcal{F}}_t(\tilde{A}_{t-1}, V_{t-1})$  est concave linéaire par morceaux par rapport à  $V_{t-1}$  pour tout  $t = 2 \dots T$ .  $\square$

#### 5.4.2 Concavité par rapport à l'état *apport passé*

Il est fréquent dans la littérature que l'on étend le caractère concave linéaire par morceaux de la fonction de bénéfice futur  $\tilde{\mathcal{F}}_t$  à l'apport passé  $\tilde{A}_{t-1}$  ce qui nous permet de rassembler l'état *volume*  $V_{t-1}$  et l'état *apport passé*  $\tilde{A}_{t-1}$  et dire, tout simplement, que  $\tilde{\mathcal{F}}_t$  est concave linéaire par morceaux sur son domaine de définition (i.e. par rapport à l'état  $S_{t-1} = (\tilde{A}_{t-1}, V_{t-1})$ ).

Ceci est possible lorsque les apports suivent un modèle PAR(1) de la forme

$$\frac{\mathbf{A}_t - \mu_t}{\sigma_t} = \phi_t \frac{\mathbf{A}_{t-1} - \mu_{t-1}}{\sigma_{t-1}} + \epsilon_t.$$

En effet, l'équation précédente nous permet d'écrire

$$\tilde{\mathbf{A}}_t = \underbrace{\phi_t \frac{\sigma_t}{\sigma_{t-1}}}_{\substack{\text{coefficient} \\ \text{indépendant de} \\ \text{la valeur de} \\ \tilde{A}_{t-1}}} \tilde{A}_{t-1} - \underbrace{\phi_t \frac{\sigma_t}{\sigma_{t-1}} \mu_{t-1} + \mu_t + \sigma_t \epsilon_t}_{\substack{\text{terme} \\ \text{indépendant de} \\ \text{la valeur de} \\ \tilde{A}_{t-1}}}.$$

Ainsi, indirectement,  $\tilde{A}_{t-1}$  fait partie du second membre des contraintes  $\mathbf{V}_t = f(V_{t-1}, \tilde{\mathbf{A}}_t, \mathbf{X}_t)$  dans  $\max_{\substack{\mathbf{X}_t, \mathbf{V}_t \\ \mathbf{V}_t=f(V_{t-1}, \tilde{\mathbf{A}}_t, \mathbf{X}_t)}} B_t(V_{t-1}, \tilde{\mathbf{A}}_t, \mathbf{X}_t) + \tilde{\mathcal{F}}_{t+1}(\tilde{\mathbf{A}}_t, \mathbf{V}_t).$

Une preuve similaire à celle développée dans la sous-section précédente nous permet de mon-

trer que pour tout  $t = 2 \dots T$ ,  $\tilde{\mathcal{F}}_t$  est concave linéaire par morceaux par rapport à  $\tilde{A}_{t-1}$ .

Malheureusement, ceci n'est plus le cas quand on considère un modèle PAR(1) de la forme

$$\frac{\ln(\mathbf{A}_t) - \mu_t}{\sigma_t} = \phi_t \frac{\ln(\mathbf{A}_{t-1}) - \mu_{t-1}}{\sigma_{t-1}} + \epsilon_t.$$

Ceci nous oblige à définir une fonction de bénéfice futur pour chaque cas d'apport passé.

Par la suite, on note  $\tilde{\mathcal{F}}_t^{(i)}(V_{t-1}) = \tilde{\mathcal{F}}_t(\tilde{A}_{t-1}^{(i)}, V_{t-1})$ ,  $t = 3 \dots T$  et  $\tilde{\mathcal{F}}_2(V_1) = \tilde{\mathcal{F}}_2(A_1, V_1)$ .

## 5.5 Représentation des fonctions de bénéfice futur

On a vu que la fonction  $V_{t-1} \mapsto \tilde{\mathcal{F}}_t^{(i)}(V_{t-1})$  est concave linéaire par morceaux, et comme  $V_{t-1}$  est borné elle peut être exprimée comme maximum d'un nombre fini d'hyperplans supports, i.e. il existe  $K_t^{(i)} < \infty$  et  $(\alpha_{t,m}^{(i)}, \beta_{t,m}^{(i)})$   $m = 1 \dots K_t^{(i)}$  tels que pour tout  $V_{t-1}$

$$\begin{aligned} \tilde{\mathcal{F}}_t^{(i)}(V_{t-1}) &= \max_{m=1 \dots K_t^{(i)}} \left( \alpha_{t,m}^{(i)T} V_{t-1} + \beta_{t,m}^{(i)} \right) = \max_{\Theta_t^{(i)}} \Theta_t^{(i)} \\ &\text{s.à } \Theta_t^{(i)} \leq \alpha_{t,m}^{(i)T} V_{t-1} + \beta_{t,m}^{(i)} \quad m = 1 \dots K_t^{(i)}. \end{aligned}$$

Ainsi pour représenter de façon complète la fonction  $\tilde{\mathcal{F}}_t^{(i)}$  il suffit de trouver les pentes  $(\alpha_{t,m}^{(i)})_{m=1 \dots K_t^{(i)}}$  et les ordonnées à l'origine  $(\beta_{t,m}^{(i)})_{m=1 \dots K_t^{(i)}}$ .

Comment trouver alors l'expression ainsi décrite de  $\tilde{\mathcal{F}}_t^{(i)}$  ?

Supposons que l'on dispose de  $\tilde{\mathcal{F}}_{t+1}^{(j)}$  pour un certain  $t \in \{2 \dots T\}$  et pour tout  $j = 1 \dots N$ , on a, en tout point  $V_{t-1}$ ,

$$\begin{aligned} \tilde{\mathcal{F}}_t^{(i)}(V_{t-1}) &= \mathbb{E}_{\tilde{\mathbf{A}}_t | \tilde{A}_{t-1}^{(i)}} \left[ \max_{\substack{\mathbf{X}_t, \mathbf{V}_t \\ \mathbf{V}_t = f(V_{t-1}, \tilde{\mathbf{A}}_t, \mathbf{X}_t)}} B_t(V_{t-1}, \tilde{\mathbf{A}}_t, \mathbf{X}_t) + \tilde{\mathcal{F}}_{t+1}(\tilde{\mathbf{A}}_t, \mathbf{V}_t) \right] \\ &= \sum_{j=1}^N P_{ij}^t \left( \max_{\substack{\mathbf{X}_t, \mathbf{V}_t \\ \mathbf{V}_t = f(V_{t-1}, \tilde{\mathbf{A}}_t^{(j)}, \mathbf{X}_t)}} B_t(V_{t-1}, \tilde{\mathbf{A}}_t^{(j)}, \mathbf{X}_t) + \tilde{\mathcal{F}}_{t+1}^{(j)}(\mathbf{V}_t) \right). \end{aligned}$$

Donc

$$\partial \tilde{\mathcal{F}}_t^{(i)}(V_{t-1}) = \sum_{j=1}^N P_{ij}^t \partial \left( \max_{\substack{\mathbf{X}_t, \mathbf{V}_t \\ \mathbf{V}_t = f(\cdot, \tilde{\mathbf{A}}_t^{(j)}, \mathbf{X}_t)}} B_t(\cdot, \tilde{\mathbf{A}}_t^{(j)}, \mathbf{X}_t) + \tilde{\mathcal{F}}_{t+1}^{(j)}(\mathbf{V}_t) \right) (V_{t-1})$$

(l'opérateur de sous-gradient  $\partial$  étant pris par rapport à la variable  $V_{t-1}$ ).

Or, pour tout  $j = 1 \dots N$

$$\begin{aligned} \max_{\substack{\mathbf{X}_t, \mathbf{V}_t \\ \mathbf{V}_t = f(V_{t-1}, \tilde{\mathbf{A}}_t^{(j)}, \mathbf{X}_t)}} B_t(V_{t-1}, \tilde{\mathbf{A}}_t^{(j)}, \mathbf{X}_t) + \tilde{\mathcal{F}}_{t+1}^{(j)}(\mathbf{V}_t) &= \max_{\mathbf{X}_t, \mathbf{V}_t} B_t(V_{t-1}, \tilde{\mathbf{A}}_t^{(j)}, \mathbf{X}_t) + \Theta_{t+1}^{(j)} \\ \text{s.à } \mathbf{V}_t &= f(V_{t-1}, \tilde{\mathbf{A}}_t^{(j)}, \mathbf{X}_t) \quad \pi_t^{(j)} \\ \Theta_{t+1}^{(j)} &\leq \alpha_{t+1,m}^{(j)T} \mathbf{V}_t + \beta_{t+1,m}^{(j)} \quad m = 1 \dots K_{t+1}^{(j)} \end{aligned}$$

où  $\pi^{(j)}$  sont des valeurs duales optimales des contraintes de conservation de masse des réservoirs à capacité positive (les contraintes dans lesquelles  $V_{t-1}$  apparait dans le second membre).

Donc

$$\partial \tilde{\mathcal{F}}_t^{(i)}(V_{t-1}) = \sum_{j=1}^N P_{ij}^t \pi_t^{(j)}.$$

Ainsi on obtient l'hyperplan support  $\mathcal{H}$  de  $\tilde{\mathcal{F}}_t^{(i)}$  au point  $V_{t-1}$

$$\mathcal{H} : \Theta_t^{(i)} \leq \partial \tilde{\mathcal{F}}_t^{(i)}(V_{t-1})^T (\mathbf{V}_{t-1} - V_{t-1}) + \tilde{\mathcal{F}}_t^{(i)}(V_{t-1})$$

Soit

$$\mathcal{H} : \Theta_t^{(i)} \leq \underbrace{\partial \tilde{\mathcal{F}}_t^{(i)}(V_{t-1})^T}_{\alpha_t^{(i)T}} \mathbf{V}_{t-1} + \underbrace{\tilde{\mathcal{F}}_t^{(i)}(V_{t-1}) - \partial \tilde{\mathcal{F}}_t^{(i)}(V_{t-1})^T V_{t-1}}_{\beta_t^{(i)}}$$

Cette procédure de construction d'hyperplans supports est schématisée dans la figure 5.2.

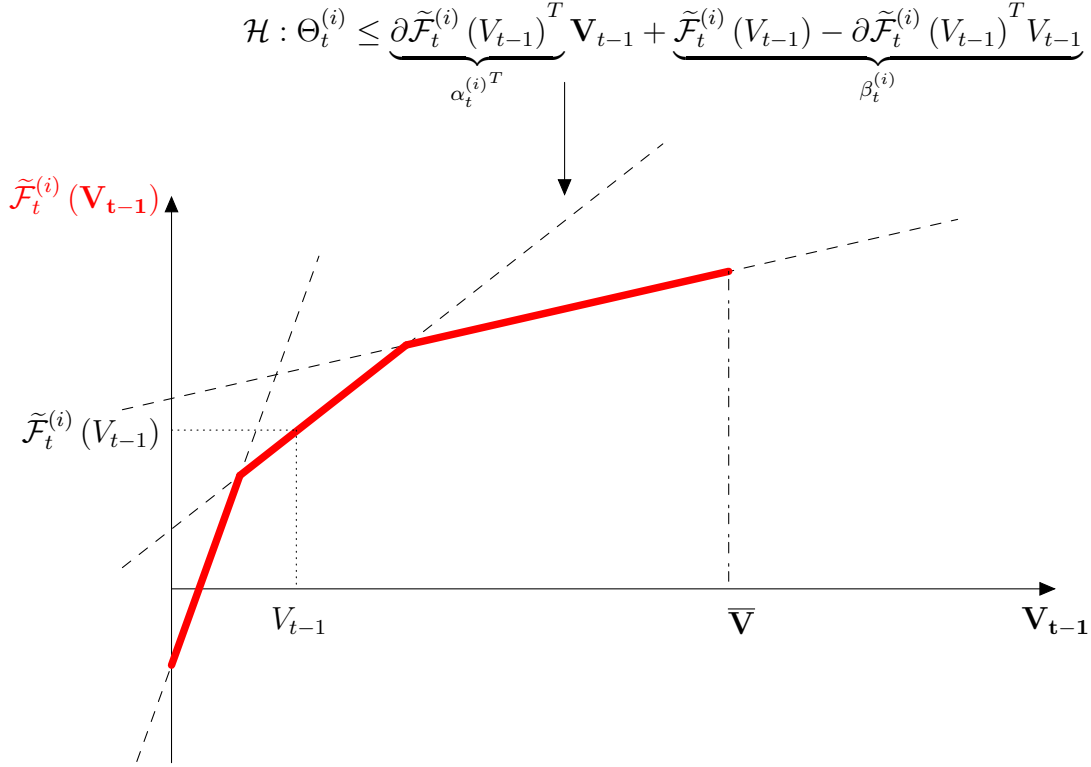


Figure 5.2 Illustration d'une fonction de bénéfice futur et ses hyperplans supports

En pratique, ce n'est pas une tâche facile de déterminer tous les hyperplans supports des fonctions  $\tilde{\mathcal{F}}_t^{(i)}$ . On se contente plutôt de les approximer par des fonctions  $\tilde{\mathfrak{F}}_t^{(i)}$  concave linéaires par morceaux dont les hyperplans supports sont des hyperplans coupants des fonctions  $\tilde{\mathcal{F}}_2$  respectives. En particulier,  $\tilde{\mathcal{F}}_2 \leq \tilde{\mathfrak{F}}_t^{(i)}$  et  $\tilde{\mathcal{F}}_t^{(i)} \leq \tilde{\mathfrak{F}}_t^{(i)}$  pour tout  $t = 3 \dots T$  et tout  $i = 1 \dots N$ . Les fonctions  $\tilde{\mathfrak{F}}_t^{(i)}$  sont définies ainsi pour tout  $t = 3 \dots T$  et tout  $i = 1 \dots N$

$$\begin{aligned} \tilde{\mathfrak{F}}_t^{(i)}(V_{t-1}) &= \mathbb{E}_{\tilde{\mathbf{A}}_t | \tilde{A}_{t-1}^{(i)}} \left[ \max_{\substack{\mathbf{X}_t, \mathbf{V}_t \\ \mathbf{V}_t = f(V_{t-1}, \tilde{\mathbf{A}}_t, \mathbf{X}_t)}} B_t(V_{t-1}, \tilde{\mathbf{A}}_t, \mathbf{X}_t) + \tilde{\mathfrak{F}}_{t+1}^{(i)}(\tilde{\mathbf{A}}_t, \mathbf{V}_t) \right] \\ &= \sum_{j=1}^N P_{ij}^t \left( \max_{\substack{\mathbf{X}_t, \mathbf{V}_t \\ \mathbf{V}_t = f(V_{t-1}, \tilde{\mathbf{A}}_t^{(j)}, \mathbf{X}_t)}} B_t(V_{t-1}, \tilde{\mathbf{A}}_t^{(j)}, \mathbf{X}_t) + \tilde{\mathfrak{F}}_{t+1}^{(j)}(\mathbf{V}_t) \right) \end{aligned}$$

avec  $\tilde{\mathfrak{F}}_{T+1}^{(j)} \equiv 0 \quad j = 1 \dots N$ .

On définit, de même, l'approximation  $\tilde{\mathfrak{F}}_2$  de  $\tilde{\mathcal{F}}_2$

$$\begin{aligned}\tilde{\mathfrak{F}}_2(V_1) &= \mathbb{E}_{\tilde{\mathbf{A}}_2|A_1} \left[ \max_{\substack{\mathbf{X}_2, \mathbf{V}_2 \\ \mathbf{V}_2 = f(V_1, \tilde{\mathbf{A}}_2, \mathbf{X}_2)}} B_t(V_1, \tilde{\mathbf{A}}_2, \mathbf{X}_2) + \tilde{\mathfrak{F}}_3(\tilde{\mathbf{A}}_2, \mathbf{V}_2) \right] \\ &= \sum_{j=1}^N P_{ij}^2 \left( \max_{\substack{\mathbf{X}_2, \mathbf{V}_2 \\ \mathbf{V}_2 = f(V_1, \tilde{\mathbf{A}}_2^{(j)}, \mathbf{X}_2)}} B_t(V_1, \tilde{\mathbf{A}}_2^{(j)}, \mathbf{X}_2) + \tilde{\mathfrak{F}}_3^{(j)}(\mathbf{V}_2) \right).\end{aligned}$$

La valeur de l'objectif dans (5.2) est alors approximée par

$$\tilde{\Theta}_1 \approx \max_{\substack{\mathbf{X}_1, \mathbf{V}_1 \\ \mathbf{V}_1 = f(V_0, A_1, \mathbf{X}_1)}} B_1(V_0, A_1, \mathbf{X}_1) + \tilde{\mathfrak{F}}_2(\mathbf{V}_1)$$

De façon similaire, sachant les approximations  $\tilde{\mathfrak{F}}_{t+1}^{(j)}$  des fonctions  $\tilde{\mathcal{F}}_{t+1}^{(j)}$ ,  $j = 1 \dots N$ , on peut ajouter un hyperplan coupant de  $\tilde{\mathcal{F}}_t^{(i)}$  à la définition de  $\tilde{\mathfrak{F}}_t^{(i)}$ .

La figure 5.3 illustre les deux fonctions :  $\tilde{\mathcal{F}}_t^{(i)}$  et son approximation  $\tilde{\mathfrak{F}}_t^{(i)}$

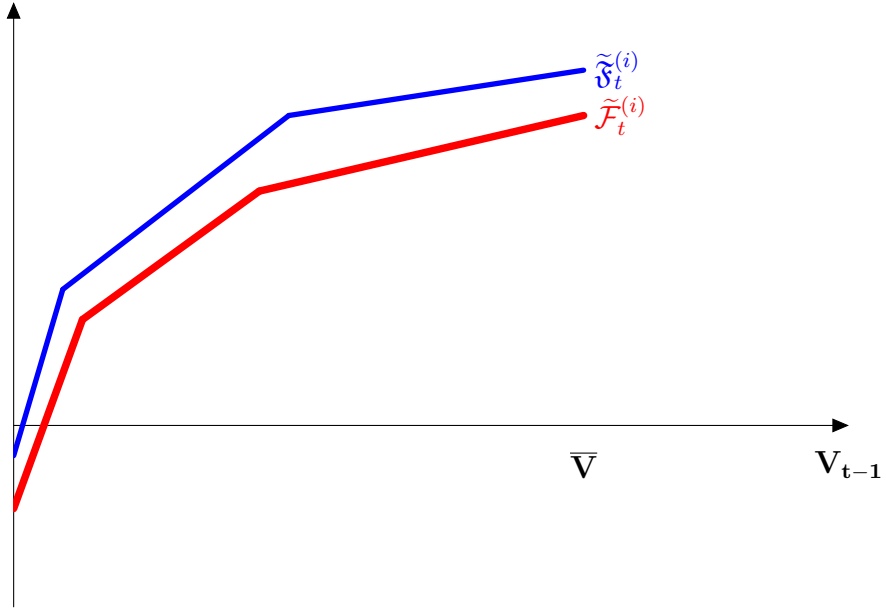


Figure 5.3 Approximation d'une fonction de bénéfice futur par des hyperplans coupants

## 5.6 Méthode de programmation dynamique stochastique exacte

Le problème (5.2) peut-être résolu à l'optimalité. Des approximations  $\tilde{\mathfrak{F}}_t^{(i)}$  (respectivement  $\tilde{\mathfrak{F}}_2$ ) aux fonctions  $\tilde{\mathcal{F}}_t^{(i)}$  (respectivement à la fonction  $\tilde{\mathcal{F}}_2$ ) sont construites récursivement de  $t = T$  à  $t = 2$  par un algorithme de programmation dynamique.

Plus précisément, étant données des approximations  $\tilde{\mathfrak{F}}_{t+1}^{(j)}$  aux fonctions  $\tilde{\mathcal{F}}_{t+1}^{(j)}$   $j = 1 \dots N$  et une discrétisation  $\mathcal{D}$  de l'espace *volume*  $\mathcal{V} = \Pi_{r \in \mathcal{R}^+} [0, \bar{v}_r]$ , on construit un hyperplan coupant de  $\tilde{\mathcal{F}}_t^{(i)}$  en tout point de  $\mathcal{D}$  et on l'ajoute à la définition de  $\tilde{\mathfrak{F}}_t^{(i)}$ .

Plus la discrétisation  $\mathcal{D}$  est exhaustive, plus les approximations  $\tilde{\mathfrak{F}}_t^{(i)}$  s'approchent des vraies fonctions  $\tilde{\mathcal{F}}_t^{(i)}$ . Rigoureusement, étant donné que l'espace *volume*  $\mathcal{V}$  est borné, le nombre d'hyperplans supports de chaque fonction  $\tilde{\mathcal{F}}_t^{(i)}$  est fini, on peut alors affirmer qu'à partir d'une valeur assez grande de  $|\mathcal{D}|$ , on aura

$$\begin{cases} \tilde{\mathfrak{F}}_2 = \tilde{\mathcal{F}}_2 \\ \tilde{\mathfrak{F}}_t^{(i)} = \tilde{\mathcal{F}}_t^{(i)} \quad \forall t \in \{3, \dots, T\} \quad \forall i \in \{1, \dots, N\}. \end{cases}$$

L'algorithme suivant permet de construire récursivement les fonctions  $\tilde{\mathfrak{F}}_t^{(i)}$ .



---

**ALGORITHME DE PROGRAMMATION DYNAMIQUE STOCHASTIQUE**


---

- **Données :**  $\mathcal{D}$  discrétisation de l'espace *volume*  $\left(\mathcal{D} \subset \mathbb{R}^{|\mathcal{R}^+|}\right)$
- **Pour tout**  $t = T \dots 2$ 
  - **Pour tout**  $V_{t-1} \in \mathcal{D}$ 
    1. **Pour tout**  $j = 1 \dots N$ 
      - (a) Résoudre  $\left(PL_t^{(j)}\right)$  : 
$$\max_{\mathbf{X}_t, \mathbf{V}_t} B_t \left(V_{t-1}, \widetilde{\mathbf{A}}_t^{(j)}, \mathbf{X}_t\right) + \widetilde{\mathfrak{F}}_{t+1}^{(j)}(\mathbf{V}_t)$$
  
 $\mathbf{V}_t = f\left(V_{t-1}, \widetilde{\mathbf{A}}_t^{(j)}, \mathbf{X}_t\right)$
      - (b) Récupérer la valeur optimale de l'objectif  $\mathcal{Z}_t^{(j)}$  de  $\left(PL_t^{(j)}\right)$  ainsi que les valeurs duales optimales  $\pi_t^{(j)}$  relatives aux contraintes de conservation de masse des réservoirs à capacité positive

2. **Si**  $t > 2$

(a) **Pour tout**  $i = 1 \dots N$

i. Calculer  $\widetilde{\mathfrak{F}}_t^{(i)}(V_{t-1}) = \sum_{j=1}^N P_{ij}^t \mathcal{Z}_t^{(j)}$

ii. Calculer  $\partial \widetilde{\mathfrak{F}}_t^{(i)}(V_{t-1}) = \sum_{j=1}^N P_{ij}^t \pi_t^{(j)}$

iii. Ajouter à la définition de  $\widetilde{\mathfrak{F}}_t^{(i)}$  l'hyperplan coupant

$$\mathcal{H} : \Theta_t^{(i)} \leq \partial \widetilde{\mathfrak{F}}_t^{(i)}(V_{t-1})^T (\mathbf{V}_{t-1} - V_{t-1}) + \widetilde{\mathfrak{F}}_t^{(i)}(V_{t-1})$$

3. **Sinon** ( $t = 2$ )

(a) Calculer  $\widetilde{\mathfrak{F}}_2(V_1) = \sum_{j=1}^N P_{i1j}^2 \mathcal{Z}_2^{(j)}$

(b) Calculer  $\partial \widetilde{\mathfrak{F}}_2(V_1) = \sum_{j=1}^N P_{i1j}^2 \pi_2^{(j)}$

(c) Ajouter à la définition de  $\widetilde{\mathfrak{F}}_2$  l'hyperplan coupant

$$\mathcal{H} : \Theta_2 \leq \partial \widetilde{\mathfrak{F}}_2(V_1)^T (\mathbf{V}_1 - V_1) + \widetilde{\mathfrak{F}}_2(V_1)$$


---

Cette méthode, quoique exacte, a une complexité exponentielle. En effet, sans restreindre la généralité, supposons que l'on discrétise le volume d'un réservoir  $r$  en  $M$  valeurs distinctes, ce qui donne  $|\mathcal{D}| = M^{|\mathcal{R}^+|}$ . En supposant que la complexité de 1.(a). est en  $\mathcal{O}(1)$ , la complexité

de la méthode ci-dessus, en d'autres termes le nombre de programmes linéaires à résoudre, est en  $\mathcal{O}\left(T \times N \times M^{|\mathcal{R}^+|}\right)$ . Le tableau 5.3 donne un ordre de grandeur de cette complexité pour  $N = 5$ ,  $M = 20$ ,  $T = 52$  (horizon d'une année) et différentes valeurs de  $|\mathcal{R}^+|$ .

Tableau 5.3 Complexité du schéma SDP exact

$ \mathcal{R}^+ $	$T \times N \times M^{ \mathcal{R}^+ }$
1	5 200
2	104 000
3	2 080 000
4	41 600 000
5	832 000 000
6	16 640 000 000

On observe bien que cette méthode serait très couteuse même pour une valeur modeste de  $M = 20$ .

### 5.7 Méthode de programmation dynamique duale stochastique (SDDP)

La méthode de programmation dynamique duale stochastique (« Stochastic Dual Dynamic Programming », en abrégé SDDP) est une généralisation de la décomposition de Benders (Benders (1962)) multi-étapes au cas d'un programme linéaire stochastique.

La décomposition de Benders multi-étapes permet de résoudre un problème linéaire déterministe dans lequel les décisions sont prises par étapes, i.e. la matrice des contraintes est par bloc mais avec des contraintes liantes qui lient les variables de décisions de deux étapes consécutives.

SDDP peut être vue comme variante de SDP mais au lieu de faire un échantillonnage *a priori* (la discrétisation  $\mathcal{D}$  utilisée précédemment), elle effectue un échantillonnage par simulations.

Comme la décomposition de Benders, SDDP est une méthode itérative dans laquelle une passe *vers l'avant* (*forward*) et une passe *vers l'arrière* (*backward*) sont effectuées à chaque itération.

### 5.7.1 Passe vers l'avant

La passe *vers l'avant* permet de trouver des états candidats  $V_1, (A_{t-1}^{(i_{t-1})}, V_{t-1})_{t=3 \dots T}$  en simulant un scénario  $(A_t^{(i_t)})_{t=2 \dots T}$  avec la politique  $\tilde{\mathfrak{F}}_2, (\tilde{\mathfrak{F}}_t^{(i)})_{t,i}$  courante. Ces états candidats seront utilisés dans la passe *vers l'arrière* pour raffiner la politique  $\tilde{\mathfrak{F}}_2, (\tilde{\mathfrak{F}}_t^{(i)})_{t,i}$ .

Le pseudo-code suivant constitue l'algorithme de la passe *vers l'avant* de SDDP.

---

#### SDDP : PASSE VERS L'AVANT

---

- **Données :** Scénario  $(\tilde{A}_t^{(i_t)})_{t=2 \dots T}$  tiré aléatoirement du processus stochastique discrétisé
- Résoudre

$$(PL_1) : \max_{\substack{\mathbf{X}_1, \mathbf{V}_1 \\ \mathbf{V}_1 = f(V_0, A_1, \mathbf{X}_1)}} B_1(V_0, A_1, \mathbf{X}_1) + \tilde{\mathfrak{F}}_2(\mathbf{V}_1)$$

- Conserver la valeur optimal  $V_1 = \mathbf{V}_1$  (volume initial pour l'étape suivante)
- Conserver l'état candidat  $V_1$  pour la passe *vers l'arrière*
- **Pour tout**  $t = 2 \dots T$

1. Résoudre  $(PL_t^{(i_t)}) :$ 

$$\max_{\substack{\mathbf{X}_t, \mathbf{V}_t \\ \mathbf{V}_t = f(V_{t-1}, \tilde{A}_t^{(i_t)}, \mathbf{X}_t)}} B_t(V_{t-1}, \tilde{A}_t^{(i_t)}, \mathbf{X}_t) + \tilde{\mathfrak{F}}_{t+1}^{(i_t)}(\mathbf{V}_t)$$

2. **Si**  $t < T$

- (a) Conserver la valeur optimal  $V_t = \mathbf{V}_t$  (volume initial pour l'étape suivante)
  - (b) Conserver l'état candidat  $(\tilde{A}_t^{(i_t)}, V_t)$  pour la passe *vers l'arrière*
- 

### 5.7.2 Passe vers l'arrière

La passe *vers l'arrière* permet de raffiner les approximations  $\tilde{\mathfrak{F}}_2, \tilde{\mathfrak{F}}_t^{(i_{t-1})}$  aux candidats respectifs  $V_1, (\tilde{A}_{t-1}^{(i_{t-1})}, V_{t-1})_{t=3 \dots T}$  issus de la passe *vers l'avant*. Un hyperplan coupant (ou coupe de Benders) est ajouté à  $\tilde{\mathfrak{F}}_t^{(i_{t-1})}$  au point  $V_{t-1}$  pour  $t = T \dots 3$ , et ensuite, de même, un hyperplan coupant (ou coupe de Benders) est ajouté à  $\tilde{\mathfrak{F}}_2$  au point  $V_1$ .

Le pseudo-code suivant décrit l'algorithme de la passe *vers l'arrière* de SDDP.

---

**SDDP : PASSE VERS L'ARRIÈRE**


---

- **Données :** États candidats  $V_1, (\tilde{A}_{t-1}^{(i_{t-1})}, V_{t-1})_{t=3 \dots T}$
- **Pour tout**  $t = T \dots 2$

1. **Pour tout**  $j = 1 \dots N$

- (a) Résoudre  $(PL_t^{(j)})$  : 
$$\max_{\mathbf{x}_t, \mathbf{V}_t} B_t(V_{t-1}, \tilde{\mathbf{A}}_t^{(j)}, \mathbf{x}_t) + \tilde{\mathfrak{F}}_{t+1}^{(j)}(\mathbf{V}_t)$$
  

$$\mathbf{V}_t = f(V_{t-1}, \tilde{\mathbf{A}}_t^{(j)}, \mathbf{x}_t)$$
- (b) Récupérer la valeur optimale de l'objectif  $\mathcal{Z}_t^{(j)}$  de  $(PL_t^{(j)})$  ainsi que les valeurs duales optimales  $\pi_t^{(j)}$  relatives aux contraintes de conservation de masse des réservoirs à capacité positive

2. **Si**  $t > 2$

- (a) Calculer  $\tilde{\mathfrak{F}}_t^{(i_{t-1})}(V_{t-1}) = \sum_{j=1}^N P_{i_{t-1}j}^t \mathcal{Z}_t^{(j)}$
- (b) Calculer  $\partial \tilde{\mathfrak{F}}_t^{(i_{t-1})}(V_{t-1}) = \sum_{j=1}^N P_{i_{t-1}j}^t \pi_t^{(j)}$
- (c) Ajouter à la définition de  $\tilde{\mathfrak{F}}_t^{(i_{t-1})}$  l'hyperplan coupant

$$\mathcal{H} : \Theta_t^{(i_{t-1})} \leq \partial \tilde{\mathfrak{F}}_t^{(i_{t-1})}(V_{t-1})^T (\mathbf{V}_{t-1} - V_{t-1}) + \tilde{\mathfrak{F}}_t^{(i_{t-1})}(V_{t-1})$$

3. **Sinon** ( $t = 2$ )

- (a) Calculer  $\tilde{\mathfrak{F}}_2(V_1) = \sum_{j=1}^N P_{i_1j}^2 \mathcal{Z}_2^{(j)}$
- (b) Calculer  $\partial \tilde{\mathfrak{F}}_2(V_1) = \sum_{j=1}^N P_{i_1j}^2 \pi_2^{(j)}$
- (c) Ajouter à la définition de  $\tilde{\mathfrak{F}}_2$  l'hyperplan coupant

$$\mathcal{H} : \Theta_2 \leq \partial \tilde{\mathfrak{F}}_2(V_1)^T (\mathbf{V}_1 - V_1) + \tilde{\mathfrak{F}}_2(V_1)$$


---

### 5.7.3 L'algorithme SDDP

Comme mentionné avant, la méthode SDDP effectue une passe *vers l'avant* suivie d'une passe *vers l'arrière* à chaque itération. Souvent, on fixe un nombre maximum d'itérations  $I_{max}$ . Ci-après le pseudo-code de l'algorithme SDDP.

---

**SDDP : L'ALGORITHME**


---

- **Données :** Un nombre maximum d'itérations  $I_{max}$
  - **Pour tout**  $l = 1 \dots I_{max}$ 
    1. Exécuter une PASSE *VERS L'AVANT*
    2. Exécuter une PASSE *VERS L'ARRIÈRE*
- 

**5.7.4 Critère d'arrêt**

Un des critères d'arrêt les plus utilisés dans SDDP est le critère d'intervalle de confiance : une borne inférieure et un intervalle de confiance sur la borne supérieure de  $\tilde{\Theta}$  (5.2) sont calculés. Le critère d'arrêt consiste à tester si la borne inférieure est à l'intérieur de l'intervalle de confiance.

**Borne supérieure**

Soient  $(X_1^*, V_1^*)$  tel que

$$(X_1^*, V_1^*) \in \underset{\substack{\mathbf{X}_1, \mathbf{V}_1 \\ \mathbf{V}_1 = f(V_0, A_1, \mathbf{X}_1)}}{\operatorname{argmax}} B_1(V_0, A_1, \mathbf{X}_1) + \tilde{\mathfrak{F}}_2(\mathbf{V}_1) \quad (5.5)$$

et soient  $(X_1^{**}, V_1^{**})$  tels que

$$(X_1^{**}, V_1^{**}) \in \underset{\substack{\mathbf{X}_1, \mathbf{V}_1 \\ \mathbf{V}_1 = f(V_0, A_1, \mathbf{X}_1)}}{\operatorname{argmax}} B_1(V_0, A_1, \mathbf{X}_1) + \tilde{\mathcal{F}}_2(\mathbf{V}_1). \quad (5.6)$$

(5.5) implique

$$B_1(V_0, A_1, X_1^*) + \tilde{\mathfrak{F}}_2(X_1^*) \geq B_1(V_0, A_1, X_1^{**}) + \tilde{\mathfrak{F}}_2(X_1^{**}).$$

Or  $\tilde{\mathfrak{F}}_2$  est une sur-approximation de  $\tilde{\mathcal{F}}_2$ , donc

$$B_1(V_0, A_1, X_1^{**}) + \tilde{\mathfrak{F}}_2(X_1^{**}) \geq B_1(V_0, A_1, X_1^{**}) + \tilde{\mathcal{F}}_2(X_1^{**}).$$

Les deux inégalités précédentes donnent

$$B_1(V_0, A_1, X_1^*) + \tilde{\mathfrak{F}}_2(X_1^*) \geq B_1(V_0, A_1, X_1^{**}) + \tilde{\mathcal{F}}_2(X_1^{**}).$$

Ainsi

$$\max_{\substack{\mathbf{X}_1, \mathbf{V}_1 \\ \mathbf{V}_1 = f(V_0, A_1, \mathbf{X}_1)}} B_1(V_0, A_1, \mathbf{X}_1) + \tilde{\mathfrak{F}}_2(\mathbf{V}_1) \geq \tilde{\Theta}_1. \quad (5.7)$$

On note la borne supérieure

$$\bar{\Theta}_1 = \max_{\substack{\mathbf{X}_1, \mathbf{V}_1 \\ \mathbf{V}_1 = f(V_0, A_1, \mathbf{X}_1)}} B_1(V_0, A_1, \mathbf{X}_1) + \tilde{\mathfrak{F}}_2(\mathbf{V}_1).$$

### Borne inférieure

Notons  $S$  le nombre de scénarios  $(\tilde{\mathbf{A}}_t)_t$  possibles ( $S = \mathcal{O}(N^T)$ ).

Soient  $(Z_s^{**})_s$  les valeurs optimales de l'objectif si l'on simule les  $S$  scénarios avec les fonctions de bénéfices futurs  $\tilde{\mathcal{F}}_2$  et  $(\tilde{\mathcal{F}}_t^{(i)})_{t,i}$ . (5.4) peut être réécrite

$$\tilde{\Theta}_1 = \sum_{s=1}^S P(s) Z_s^{**} \quad (5.8)$$

où  $P(s)$  est la probabilité du scénario  $s$ .

Soient  $(Z_s^*)_s$  les valeurs optimales de l'objectif si l'on simule les  $S$  scénarios avec les fonctions de bénéfices futurs approximées  $\tilde{\mathfrak{F}}_2$  et  $(\tilde{\mathfrak{F}}_t^{(i)})_{t,i}$ . Puisque ces valeurs optimales ont été obtenues en utilisant une politique sous-optimale, on a l'inégalité suivante

$$\sum_{s=1}^S P(s) Z_s^* \leq \sum_{s=1}^S P(s) Z_s^{**}, \quad (5.9)$$

soit

$$\sum_{s=1}^S P(s) Z_s^* \leq \tilde{\Theta}_1. \quad (5.10)$$

### Estimateur de la borne inférieure

Le calcul de la borne inférieure en simulant l'ensemble de tous les scénarios possibles demande beaucoup de temps de calcul. En pratique, on tire un nombre  $S'$  de scénarios ( $S' \ll S$ ) et on calcule une estimation statistique de la borne inférieure (5.10)

$$\frac{1}{S'} \sum_{s=1}^{S'} Z_s^* \approx \sum_{s=1}^S P(s) Z_s^*. \quad (5.11)$$

Notons que pour que la dernière inégalité soit juste il faut que le tirage aléatoire des  $S'$  scénarios soit effectué avec remise.

### Intervalle de confiance

Notons  $\underline{\Theta}_1 = \frac{1}{S'} \sum_{s=1}^{S'} Z_s^*$  et soit l'écart-type empirique  $\sigma = \sqrt{\frac{1}{S'-1} \sum_{s=1}^{S'} (Z_s^* - \underline{\Theta}_1)^2}$

$\left[ \underline{\Theta}_1 - z_{\alpha/2} \frac{\sigma}{\sqrt{S'}}, \underline{\Theta}_1 + z_{\alpha/2} \frac{\sigma}{\sqrt{S'}} \right]$  est un intervalle de confiance à  $100(1 - \alpha)\%$  pour  $\Theta$  où  $z_{\alpha/2}$  est le  $(1 - \alpha)$ -quantile de la loi normale centrée réduite. On prend souvent  $\alpha = 0.5$  ce qui donne  $z_{0.025} = 1.96$ .

### 5.7.5 Convergence

Deux articles majeurs prouvent la convergence de la méthode SDDP : l'article de Philpott et Guan (2008) et celui de Shapiro (2011). La preuve de convergence dans Philpott et Guan (2008) se base sur la preuve de convergence de la décomposition de Benders multi-étapes en la généralisant au cas stochastique. Il démontre que le nombre d'hyperplans coupants que l'on peut générer pour définir chaque fonction de bénéfice futur  $\tilde{\mathfrak{F}}_t^{(i)}$  (ou  $\tilde{\mathfrak{F}}_2$ ) est fini. Plus encore, il démontre que, sous certaines hypothèses non restrictives, les approximations  $\tilde{\mathfrak{F}}_2, \left(\tilde{\mathfrak{F}}_t^{(i)}\right)_{t,i}$  convergent vers les vraies fonctions  $\tilde{\mathcal{F}}_2, \left(\tilde{\mathcal{F}}_t^{(i)}\right)_{t,i}$  après un nombre fini d'itérations.

Les hypothèses en question sont :

- les ensembles réalisables des programmes linéaires à une étape sont bornés et non vides, ce qui est vérifié dans notre cas car le volume de chaque réservoir est borné et on pénalise la demande non-satisfaite ;
- chaque scénario possible (i.e. dont la probabilité d'occurrence est positive) du processus stochastique  $\left(\tilde{\mathbf{A}}_t\right)_t$  doit pouvoir être tiré une infinité de fois dans la passe *vers l'avant*, cette hypothèse est vérifiée dans notre cas si l'on utilise un vrai générateur de nombres aléatoires de loi  $\mathcal{U}(0, 1)$ . En pratique, on utilise un pseudo-générateur de nombres aléatoires ;
- chaque réalisation possible de la variable aléatoire  $\tilde{\mathbf{A}}_t$  doit pouvoir être tirée une infinité de fois lors de l'ajout de coupes à une fonction  $\tilde{\mathfrak{F}}_t^{(i)}$  dans la passe *vers l'arrière*. Cette hypothèse n'est pas nécessaire dans notre cas car on utilise toutes les  $N$  réalisations possibles pour calculer une nouvelle coupe. La preuve de Philpott et Guan (2008) est plus générale dans le sens qu'elle n'exige pas d'utiliser toutes les réalisations possibles (si  $N$  est grand cela peut-être fastidieux).

## 5.8 Méthode de programmation dynamique stochastique approchée (ASDP)

Cette section est basée sur l'article en cours de rédaction intitulé « Adaptive discretization of the state space for stochastic dynamic programming applied to multireservoir system » des auteurs Stéphane Krau, Grégory Émiel, James Merleau et Robert Leconte. On essaiera dans cette section de présenter l'idée générale sans rentrer dans les petits détails qui sont décrits dans l'article précité.

La méthode ASDP est très similaire à la méthode SDP présenté précédemment. La seule

différence consiste à utiliser une discrétisation adaptée de l'espace des états  $\mathcal{D}_{t-1}^{(i)}$  au lieu de la discrétisation *a priori*  $\mathcal{D}$  pour construire la fonction de bénéfice futur  $\tilde{\mathfrak{F}}_t^{(i)}$  utilisée dans SDP. Cette discrétisation  $\mathcal{D}_{t-1}^{(i)}$ , bien entendu, dépend de  $t$  et de  $i$  (pour  $t = 2$  on utilise une seule discrétisation  $\mathcal{D}_1$  puisqu'une seule fonction de bénéfice futur  $\tilde{\mathcal{F}}_2$  est à déterminer). Le but étant de ne pas discrétiser de façon uniforme l'espace des états, mais plutôt discrétiser là où la fonction de bénéfice futur  $\tilde{\mathfrak{F}}_t^{(i)}$  présente une forte courbure.

### 5.8.1 Discrétisation adaptée de l'espace d'état *volume*

Soient  $t \in \{3 \dots T\}$  et  $i \in \{1 \dots N\}$ . Notons  $n = |\mathcal{R}^+|$ , la dimension de l'espace d'état *volume*.

L'espace *volume*  $\mathcal{V} = \Pi_{r \in \mathcal{R}^+} [0, \overline{v}_r]$  et partitionné en sous-régions spéciales de forme d'hyper-rectangles. Initialement, la discrétisation  $\mathcal{D}_{t-1}^{(i)}$  de  $\mathcal{V}$  est formée des points extrêmes de  $\mathcal{V}$ , i.e.  $\mathcal{D}_{t-1}^{(i)} \leftarrow \Pi_{r \in \mathcal{R}^+} \{0, \overline{v}_r\}$ ,  $\mathcal{D}_{t-1}^{(i)}$  est un seul hyper-rectangle. L'algorithme de discrétisation adaptée consiste à découper chaque hyper-rectangle de  $\mathcal{D}_{t-1}^{(i)}$  en sous hyper-rectangles selon un critère donné.

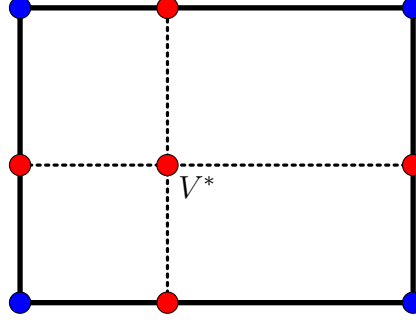
Un façon de découper un hyper-rectangle  $\mathcal{H}$  est de trouver le point intérieur de  $\mathcal{H}$  qui maximise l'erreur sur l'approximation de la fonction de bénéfice futur  $\tilde{\mathfrak{F}}_t^{(i)}$  sur  $\mathcal{H}$ . Cette erreur maximum est définie comme étant le maximum de la différence entre  $\tilde{\mathfrak{F}}_t^{(i)}$  et son interpolation sur  $\mathcal{H}$ , elle est obtenue en résolvant le programme linéaire suivant

$$\begin{aligned} \epsilon_{max}^{\mathcal{H}} = & \max_{\mathbf{V}, \boldsymbol{\lambda}} \quad \tilde{\mathfrak{F}}_t^{(i)}(\mathbf{V}) - \sum_{j=1}^n \lambda_j \tilde{\mathfrak{F}}_t^{(i)}(V_j) \\ \text{s. à} \quad & \mathbf{V} = \sum_{j=1}^n \lambda_j V_j \\ & \sum_{j=1}^n \lambda_j = 1 \\ & \boldsymbol{\lambda} \geq 0 \end{aligned} \tag{5.12}$$

où les  $V_j$  sont les  $2^n$  sommets de l'hyper-rectangle  $\mathcal{H}$ .

Soit  $V^*$  une solution optimale du programme linéaire (5.12), l'hyper-rectangle  $\mathcal{H}$  est découpé en  $2^n$  hyper-rectangles fils et  $n2^{n-1} + 1$  nouveaux points sont ajoutés à la discrétisation  $\mathcal{D}_{t-1}^{(i)}$  comme l'illustre la figure suivante



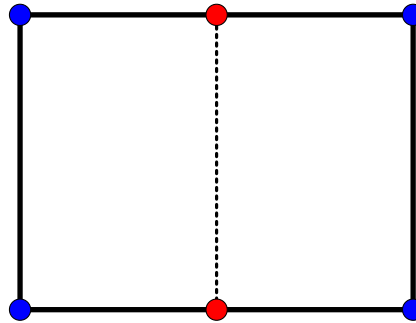


Les hyper-rectangles  $\mathcal{H}$  sont triés en temps réel par ordre décroissant de  $\varepsilon_{max}^{\mathcal{H}}$  et à chaque itération l'hyper-rectangle avec la plus grande  $\varepsilon_{max}^{\mathcal{H}}$  est découpé en  $2^n$  hyper-rectangles fils.

Cette méthode nécessite la résolution d'un programme linéaire pour chacun des  $2^n$  hyper-rectangles fils à chaque découpage d'un hyper-rectangle père, elle peut ainsi devenir couteuse à mesure que la dimension de l'espace d'état *volume* augmente. Dans la sous-section suivante, une méthode simple de découpage est présentée.

### 5.8.2 Discrétisation adaptée simplifiée de l'espace d'état *volume*

Au lieu de trouver le point intérieur d'un hyper-rectangle  $\mathcal{H}$  qui maximise l'erreur  $\varepsilon_{max}^{\mathcal{H}}$ , on calcule l'erreur sur chacune des  $n2^{n-1}$  arêtes  $e$  de  $\mathcal{H}$  et on regroupe ces erreurs par dimension, i.e. l'erreur associée à une dimension de l'espace état *volume* correspond à la somme des erreurs des  $2^{n-1}$  arêtes qui lui sont parallèles. L'hyper-rectangle est ensuite découpé en deux hyper-rectangles fils de volumes égaux suivant la dimension de plus forte erreur. Cette nouvelle manière de découper engendre seulement  $2^{n-1}$  nouveaux échantillons par hyper-rectangle découpé.



Pour éviter de découper de très petits hyper-rectangles, l'erreur associée à un hyper-rectangle est pondérée par son volume (produit des longueurs selon chaque dimension), on parle ainsi de *volume d'erreur*.

L'avantage de cette méthode de découpage est qu'elle ne nécessite pas de résolution d'un programme linéaire car l'erreur sur une arête peut être calculée analytiquement comme le montre la figure 5.4.

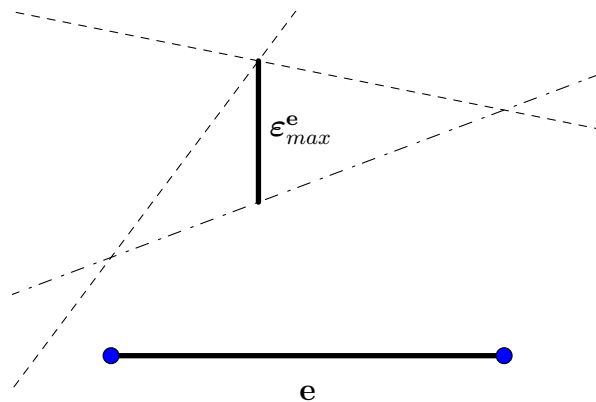


Figure 5.4 Calcul analytique de l'erreur sur une arête d'un hyper-rectangle

En effet, il suffit de trouver le point d'intersection des deux droites de l'extrapolation, l'erreur est alors la distance de ce point à la droite d'interpolation.

### 5.8.3 L'algorithme ASDP

Le pseudo-code suivant décrit l'algorithme de ASDP.

---

**ALGORITHME DE PROGRAMMATION DYNAMIQUE STOCHASTIQUE APPROCHÉE**


---

- **Données :**  $D_{max}$  nombre maximum d'échantillons de l'espace d'état *volume*
- **Pour tout**  $t = T \dots 2$ 
  - **Si**  $t > 2$ 
    - **Tant que**  $\sum_{i=1}^N |\mathcal{D}_{t-1}^{(i)}| < D_{max}$ 
      1. Découper l'hyper-rectangle  $\mathcal{H}$  de  $\cup_{i=1}^N \mathcal{D}_{t-1}^{(i)}$  ayant le plus grand volume d'erreur. Soit  $k$  l'unique élément de  $\{1 \dots N\}$  tel que  $\mathcal{H}$  fait partie de la partition de  $\mathcal{D}_{t-1}^{(k)}$
      2. **Pour tout** sommet  $V_{t-1}$  nouvellement ajouté à  $\mathcal{D}_{t-1}^{(k)}$ 
        - (a) **Pour tout**  $j = 1 \dots N$ 
          - i. Résoudre  $\left( PL_t^{(j)} \right) :$  
$$\max_{\mathbf{x}_t, \mathbf{v}_t} B_t \left( V_{t-1}, \tilde{\mathbf{A}}_t^{(j)}, \mathbf{x}_t \right) + \tilde{\mathfrak{F}}_{t+1}^{(j)} (\mathbf{v}_t)$$

$$\mathbf{v}_t = f \left( V_{t-1}, \tilde{\mathbf{A}}_t^{(j)}, \mathbf{x}_t \right)$$
          - ii. Récupérer la valeur optimale de l'objectif  $\mathcal{Z}_t^{(j)}$  de  $\left( PL_t^{(j)} \right)$  ainsi que les valeurs duales optimales  $\pi_t^{(j)}$  relatives aux contraintes de conservation de masse des réservoirs à capacité positive
        - (b) Calculer  $\tilde{\mathfrak{F}}_t^{(k)} (V_{t-1}) = \sum_{j=1}^N P_{kj}^t \mathcal{Z}_t^{(j)}$
        - (c) Calculer  $\partial \tilde{\mathfrak{F}}_t^{(k)} (V_{t-1}) = \sum_{j=1}^N P_{kj}^t \pi_t^{(j)}$
        - (d) Ajouter à la définition de  $\tilde{\mathfrak{F}}_t^{(k)}$  l'hyperplan coupant

$$\mathcal{H} : \Theta_t^{(k)} \leq \partial \tilde{\mathfrak{F}}_t^{(k)} (V_{t-1})^T (\mathbf{V}_{t-1} - V_{t-1}) + \tilde{\mathfrak{F}}_t^{(k)} (V_{t-1})$$

- **Si**  $(t = 2)$ 
  - **Tant que**  $|\mathcal{D}_1| < D_{max}$ 
    1. Découper l'hyper-rectangle  $\mathcal{H}$  de  $\mathcal{D}_1$  ayant le plus grand volume d'erreur.
    2. **Pour tout** sommet  $V_1$  nouvellement ajouté à  $\mathcal{D}_{t-1}$ 
      - (a) **Pour tout**  $j = 1 \dots N$ 
        - i. Résoudre  $\left( PL_2^{(j)} \right) :$  
$$\max_{\mathbf{x}_2, \mathbf{v}_2} B_2 \left( V_1, \tilde{\mathbf{A}}_2^{(j)}, \mathbf{x}_2 \right) + \tilde{\mathfrak{F}}_3^{(j)} (\mathbf{v}_2)$$

$$\mathbf{v}_2 = f \left( V_1, \tilde{\mathbf{A}}_2^{(j)}, \mathbf{x}_2 \right)$$
        - ii. Récupérer la valeur optimale de l'objectif  $\mathcal{Z}_2^{(j)}$  de  $\left( PL_2^{(j)} \right)$  ainsi que les valeurs duales optimales  $\pi_2^{(j)}$  relatives aux contraintes de conservation de masse des réservoirs à capacité positive
      - (b) Calculer  $\tilde{\mathfrak{F}}_2 (V_1) = \sum_{j=1}^N P_{1j}^2 \mathcal{Z}_2^{(j)}$
      - (c) Calculer  $\partial \tilde{\mathfrak{F}}_2 (V_1) = \sum_{j=1}^N P_{1j}^2 \pi_2^{(j)}$
      - (d) Ajouter à la définition de  $\tilde{\mathfrak{F}}_2$  l'hyperplan coupant

$$\mathcal{H} : \Theta_2 \leq \partial \tilde{\mathfrak{F}}_2 (V_1)^T (\mathbf{V}_1 - V_1) + \tilde{\mathfrak{F}}_2 (V_1)$$


---

#### 5.8.4 Critère d'arrêt

Au lieu de fixer *a priori* un nombre maximum  $D_{max}$  d'échantillons de l'espace d'état *volume*, on peut avoir un critère d'arrêt sur le *volume d'erreur* maximal des hyper-rectangles. Dès que le *volume d'erreur* est en bas d'une tolérance fixée, on estime que les fonctions de bénéfice futur de la semaine courante sont suffisamment approximées et on passe à la semaine précédente.

#### 5.8.5 Convergence

Tout comme SDP (cf. section 5.6), plus  $D_{max}$  est grand plus les approximations  $\tilde{\mathfrak{F}}_2, \tilde{\mathfrak{F}}_t^{(i)}$  s'approchent des vraies fonctions  $\tilde{\mathcal{F}}_t, \tilde{\mathcal{F}}_t^{(i)}$ , et on peut de même affirmer qu'à partir d'une valeur assez grande de  $D_{max}$  on aura

$$\begin{cases} \tilde{\mathfrak{F}}_2 = \tilde{\mathcal{F}}_2 \\ \tilde{\mathfrak{F}}_t^{(i)} = \tilde{\mathcal{F}}_t^{(i)} \quad \forall t \in \{3, \dots, T\} \quad \forall i \in \{1, \dots, N\} \end{cases}$$

## CHAPITRE 6 ANALYSE D'UN CAS D'ÉTUDE : LA GRANDE RIVIÈRE

Ce chapitre présente les résultats de l'application de la méthodologie vue dans les chapitres précédents sur un cas réel : La Grande rivière située au nord de la province de Québec (voir la figure 1.2). Il s'agit de la rivière qui fournit la plus importante part d'électricité au Québec.

Le chapitre est découpé en trois sections majeures. Dans la première section, on décrit le cas d'étude. Dans la deuxième section, une comparaison des deux modélisations des fonctions de production des centrales hydroélectriques vues au chapitre 3 est faite sur les centrales considérées. Enfin, on fournit dans la troisième section les résultats de l'application des méthodes SDDP et ASDP vues au chapitre 5 au problème de planification moyen-terme de la production d'hydroélectricité.

### 6.1 Description du cas d'étude

#### 6.1.1 Caractéristiques des installations hydroélectriques

LGR contient six réservoirs dont quatre sont des réservoirs à capacité positive (peuvent stocker l'eau) : Caniapiscou, La Grande<sub>4</sub>, La Grande<sub>3</sub>, La Grande<sub>2</sub>. Les réservoirs La Forge<sub>2</sub> et La Forge<sub>1</sub> sont considérés comme réservoirs sans capacité. Le tableau 6.1 donnent les caractéristiques des quatre réservoirs à capacité positive.

Tableau 6.1 Caractéristiques des réservoirs de LGR

Réservoir	Volume maximum (hm <sup>3</sup> )	Équivalent énergétique approximatif (GWh)
Caniapiscou	39 009	47 772
La Grande <sub>4</sub>	7 079	6 483
La Grande <sub>3</sub>	25 195	15 583
La Grande <sub>2</sub>	19 369	8 131

À chaque réservoir de LGR sont associés une centrale hydroélectrique et un ouvrage d'évacuation. Rappelons qu'un réservoir à capacité positive possède une centrale à hauteur de chute variable et qu'un réservoir sans capacité possède une centrale à hauteur de chute fixe. Le tableau 6.2 donne les caractéristiques des centrales de LGR, on y trouve la hauteur de chute, le débit et la puissance maximums à la hauteur de chute de référence.

Tableau 6.2 Caractéristiques des centrales de LGR

Centrale	$h_{max}$ (m)	$q_{max}^0$ (m <sup>3</sup> /s)	$p_{max}^0$ (MW)
Caniapiscou	41.34	1 133	402.90
La Forge <sub>2</sub>	-	1 192.13	280.07
La Forge <sub>1</sub>	-	1 642.90	828.67
La Grande <sub>4</sub>	119.60	2 545.75	2 656.24
La Grande <sub>3</sub>	84.20	3 552.84	2 362.84
La Grande <sub>2</sub>	145.73	6 245.97	7 394.14

Le tableau 6.3 donne les débits maximums des ouvrages d'évacuation de LGR.

Tableau 6.3 Caractéristiques des ouvrages d'évacuation de LGR

Ouvrage d'évacuation	$q_{max}$ (m <sup>3</sup> /s)
Caniapiscou	$\infty$
La Forge <sub>2</sub>	2 475.00
La Forge <sub>1</sub>	2 958.50
La Grande <sub>4</sub>	7 723.23
La Grande <sub>3</sub>	10 507.00
La Grande <sub>2</sub>	16 415.00

La figure 6.1 illustre la disposition et les relations entre les différentes composantes de LGR. On y voit en particulier que le déversement au niveau de Caniapiscou est hors-système (confluent) et il est préférable de le réduire le plus possible vu que plusieurs autres réservoirs sont en bas de Caniapiscou et pourraient donc utiliser cette eau « gaspillée ».

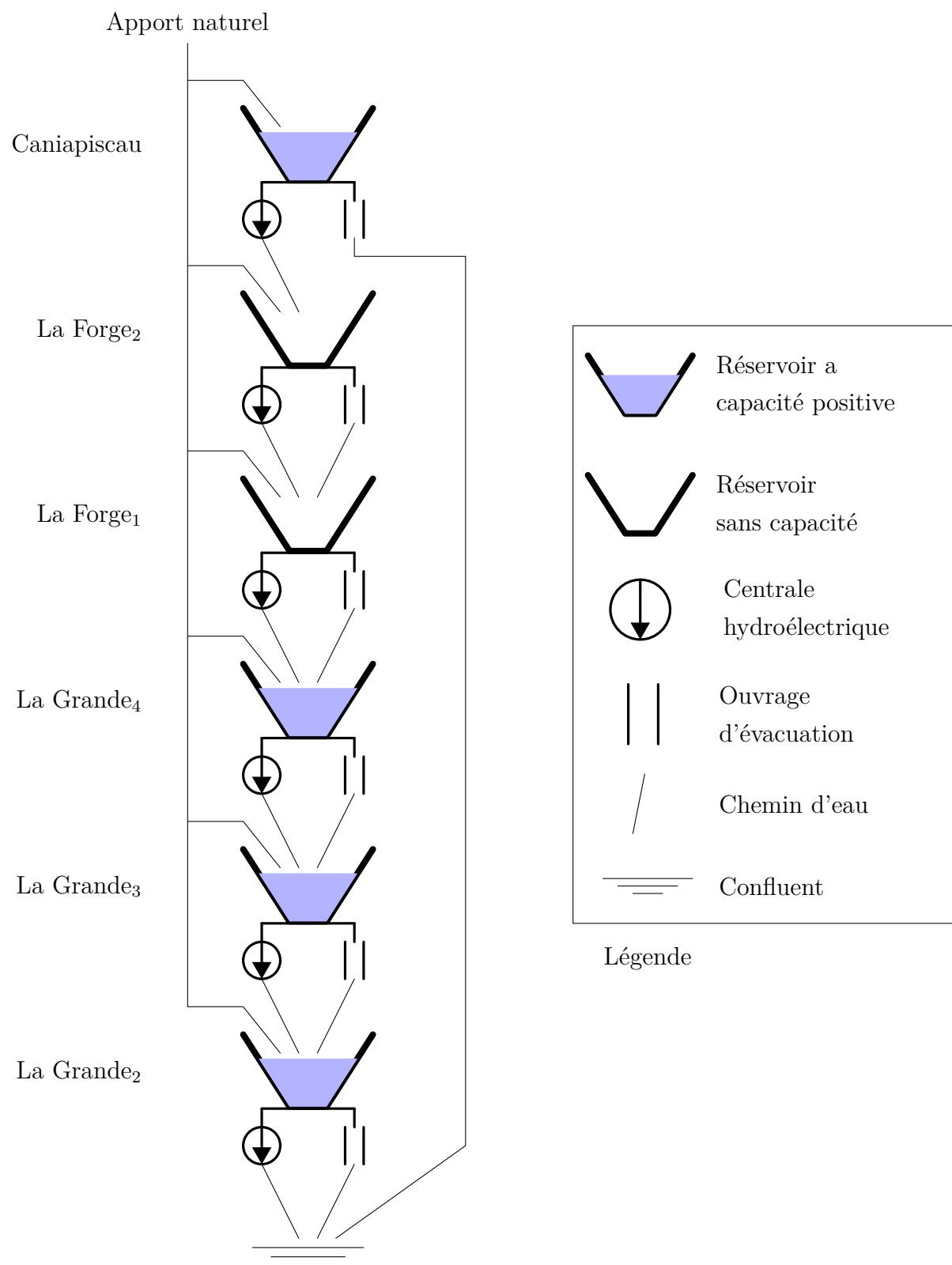


Figure 6.1 Schéma de La Grande Rivière

### 6.1.2 Demande en électricité

Les modèles de planification de la production d'hydroélectricité que l'on traite visent, entre autres, à satisfaire la demande en électricité. La figure 6.2 illustre la demande à satisfaire sur une année (on considère que la demande est stationnaire, c'est à dire qu'elle demeure inchangée d'une année à l'autre). On considère deux classe de demande : une classe pointe et une classe moyenne qui dure respectivement 42h et 126h sur une semaine (168h).



Figure 6.2 Demande électrique

## 6.2 Comparaison de différentes modélisations des fonctions de production des centrales

Chacune des trois modélisations vues au chapitre 3 sont évaluées sur une grille fine  $\mathcal{G}$  de l'espace *débit-hauteur de chute*  $(q, h)$ . Le critère d'évaluation du résultat d'une modélisation donnée est l'erreur-type sur la grille  $\mathcal{G}$ ,  $RMSE_{\mathcal{G}}$

$$RMSE_{\mathcal{G}} = \sqrt{\frac{1}{|\mathcal{G}|} \sum_{(q_i, h_i) \in \mathcal{G}} (\tilde{p}_C(q_i, h_i) - p_C(q_i, h_i))^2} \quad (6.1)$$



où  $\tilde{p}_C$  est la fonction puissance approximée ((3.8) dans le cas de **EnvConc**, (3.9) dans le cas de **PartMC** ou **PartMINMAX**) et  $p_C$  est la fonction puissance exacte (3.6).

Le tableau 6.4 donnent la taille des grilles utilisées.

Tableau 6.4 Taille des grilles utilisées

Centrale	$ \mathcal{G} $
Caniapiscau	16 962
La Grande <sub>4</sub>	32 832
La Grande <sub>3</sub>	49 521
La Grande <sub>2</sub>	56 225

### 6.2.1 Modélisation par enveloppe concave

Le tableau 6.5 donne, pour chacune des centrales à hauteur de chute variable, l'erreur-type  $RMSE_{\mathcal{G}}$  de la production estimée et le pourcentage de  $RMSE_{\mathcal{G}}$  par rapport à la puissance installée de la centrale quand on approxime les fonction de production des centrales par la méthode **EnvConc**.

Tableau 6.5 Erreur-type de la production estimée (cas de **EnvConc**)

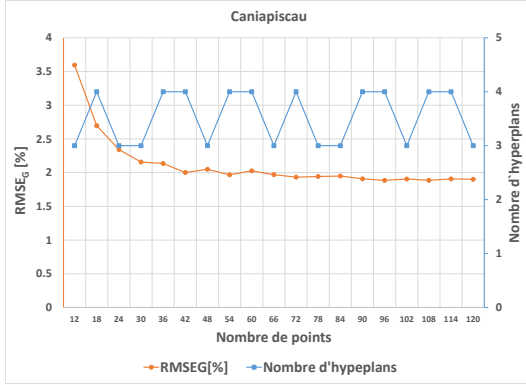
Centrale	$RMSE_{\mathcal{G}}(\text{MW})$	$RMSE_{\mathcal{G}}[\%]$
Caniapiscau	14.56	3.61
La Grande <sub>4</sub>	20.41	0.77
La Grande <sub>3</sub>	32.08	1.36
La Grande <sub>2</sub>	15.57	0.21

### 6.2.2 Modélisation par une fonction concave affine par morceaux

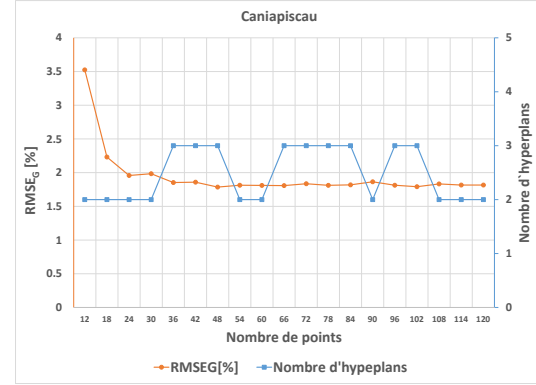
Pour les deux variantes **PartMC** et **PartMINMAX**, les plans affines sont obtenus sur un sous-ensemble de la grille constitués des points  $(q, h, p)$  de configurations optimales pour différentes valeurs de la hauteur de chute, ceci est tout à fait raisonnable car une centrale est utilisée dans l'une de ces configurations.

Les résultats de l'application de l'algorithme de partitionnement sur les fonctions de productions des 4 centrales à hauteur de chute variable du cas d'étude sont résumés dans les figures 6.3, 6.4, 6.5 et 6.6. L'erreur-type est représentée en pourcentage de la puissance installée

de la centrale. À noter que pour ces résultats, on a choisit les paramètres  $N_{essais} = 100$  et  $l_{max} = 200$ .

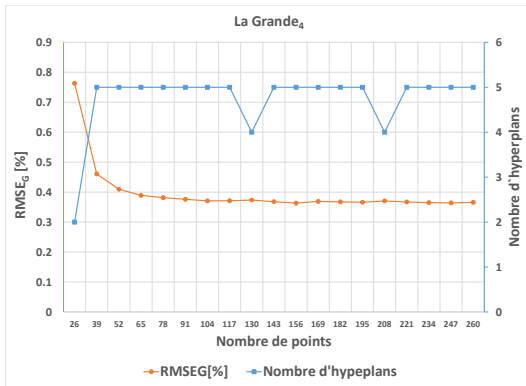


(a) **PartMC**

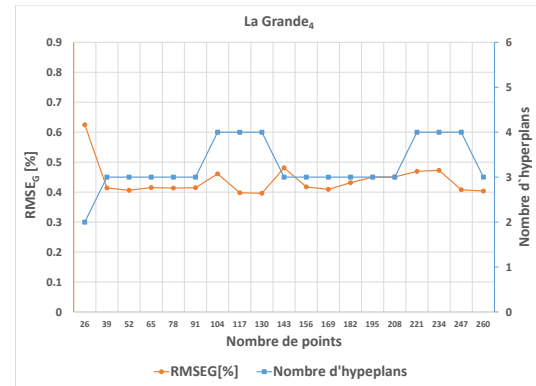


(b) **PartMINMAX**

Figure 6.3 Relation entre, (1) la discrétisation de l'espace débit-hauteur de chute, (2) l'erreur-type ( $RMSE_G$ ), et (3) le nombre d'hyperplans approximant la fonction de production (cas de la centrale Caniapiscou)



(a) **PartMC**



(b) **PartMINMAX**

Figure 6.4 Relation entre, (1) la discrétisation de l'espace débit-hauteur de chute, (2) l'erreur-type ( $RMSE_G$ ), et (3) le nombre d'hyperplans approximant la fonction de production (cas de la centrale La Grande<sub>4</sub>)

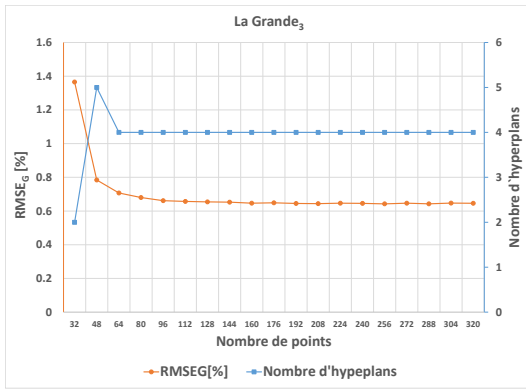
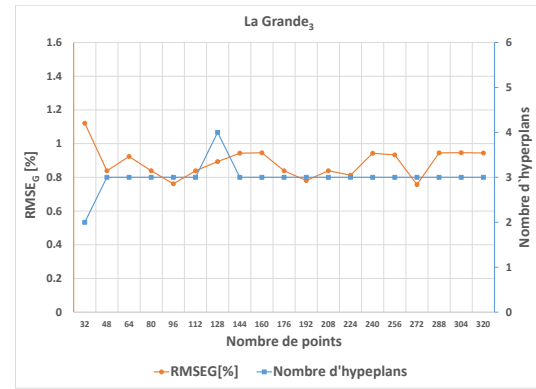
(a) **PartMC**(b) **PartMINMAX**

Figure 6.5 Relation entre, (1) la discrétisation de l'espace débit-hauteur de chute, (2) l'erreur-type ( $RMSE_G$ ), et (3) le nombre d'hyperplans approximant la fonction de production (cas de la centrale La Grande<sub>3</sub>)

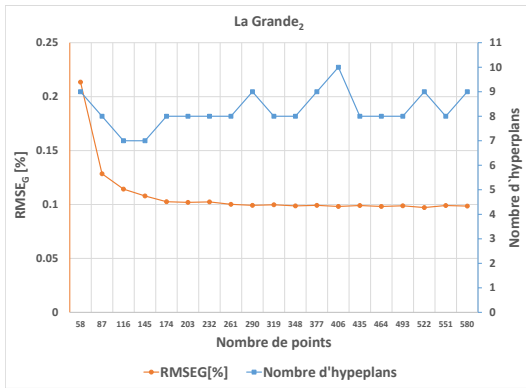
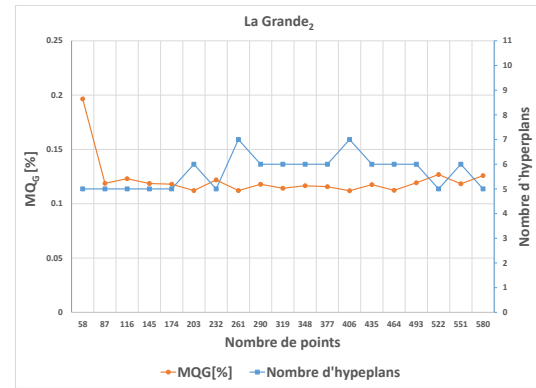
(a) **PartMC**(b) **PartMINMAX**

Figure 6.6 Relation entre, (1) la discrétisation de l'espace débit-hauteur de chute, (2) l'erreur-type ( $RMSE_G$ ), et (3) le nombre d'hyperplans approximant la fonction de production (cas de la centrale La Grande<sub>2</sub>)

Si l'on prend le cas de la centrale Caniapiscou (figure 6.3), on peut remarquer que les deux approches de partitionnement **PartMC** et **PartMINMAX** sont comparables au niveau de

l'erreur-type ( $RMSE_{\mathcal{G}}$ ). Cependant, **PartMINMAX** est meilleure que **PartMC** au niveau du nombre d'hyperplans définissant la fonction de production approximée. On peut faire le même constat pour les autres centrales.

Par ailleurs, **PartMC** est meilleure que **PartMINMAX** au niveau du temps d'exécution de l'algorithme de partitionnement (c.f. section 3.2.4 du chapitre 3). En effet, pour trouver le meilleur hyperplan au sens de l'erreur maximum la plus petite, **PartMINMAX** résout un programme linéaire assez grand, alors que **PartMC** effectue une inversion de matrice  $3 \times 3$  ou résout un problème quadratique à trois variables. Le tableau 6.6 illustre cette différence en temps d'exécution. Ici, on a pris  $N_{essais} = 100$  et  $l_{max} = 50$ .

Tableau 6.6 Ordre de grandeur du temps d'exécution de l'algorithme de partitionnement

Centrale	Nuage de points	Temps <b>PartMC</b> (s)	Temps <b>PartMINMAX</b> (s)
Caniapiscau	120	43	82
La Grande <sub>4</sub>	260	88	164
La Grande <sub>3</sub>	320	109	195
La Grande <sub>2</sub>	580	271	362

On peut voir que le temps d'exécution de l'algorithme de partitionnement avec **PartMINMAX** est à peu près le double de celui avec **PartMC**.

### 6.2.3 Conclusion

Les résultats qu'on vient d'analyser montre que les trois modélisations **EnvConc**, **PartMC** et **PartMINMAX** sont comparables et plus ou moins équivalentes en matière d'erreur-type  $RMSE_{\mathcal{G}}$ . Cependant, Elle peuvent impacter différemment la performance (temps d'exécution) des algorithmes de planification de la production d'hydroélectricité. En effet, il s'agit de modélisations complémentaires au sens de la théorie de la programmation linéaire. **EnvConc** nécessitent plusieurs variables (les  $\lambda$ ) et trois contraintes. la modélisation par fonction concave affine par morceaux (que ce soit via **PartMC** ou **PartMINMAX**) n'ajoute pas de variables mais nécessite l'ajout d'une contrainte par hyperplan.

Intuitivement, on peut faire la conjecture que la modélisation des fonctions de production par des fonctions concaves affines par morceaux est plus performante si le nombre d'hyperplans reste raisonnable, dans ce cas **PartMINMAX** semble être la plus appropriée. À la fin de ce chapitre, un retour sur les deux modélisations sera fourni.

### 6.3 Planification moyen-terme de la production d'hydroélectricité : SDDP vs ASDP

Dans cette section, on présente les résultats issus de l'application des méthodes SDDP et ASDP au problème de planification moyen-terme de la production d'hydroélectricité. L'objectif est de trouver une politique optimale de gestion de LGR sur un horizon d'une année.

On part des volumes initiaux des réservoirs listés dans le tableau 6.7.

Tableau 6.7 Volumes initiaux des réservoirs à capacité positive

Réservoir	Volume initial $V_0$
Caniapiscou	30 000
La Grande <sub>4</sub>	5 000
La Grande <sub>3</sub>	18 000
La Grande <sub>2</sub>	16 000

Quant à l'apport de la semaine 1, supposé connu à l'avance, on part d'une valeur de  $A_1 = 2418.331120 \text{ m}^3/s$ .

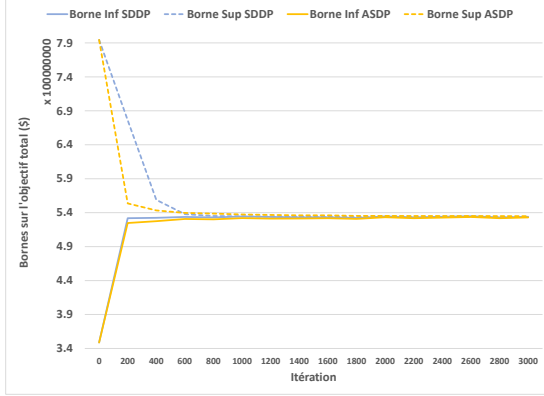
Le processus d'apport hydrique est discrétisé en  $N = 5$  valeurs pour chaque semaine.

Comme on ne connaît pas la valeur de l'eau à la fin de l'horizon considéré (fixée à  $\mathcal{F}_{T+1} \equiv 0$ ), il faut en théorie résoudre le problème de planification sur un horizon infini ( $T \rightarrow \infty$ ) ce qui n'est pas envisageable en pratique. Une façon approximative de surmonter cet obstacle est de résoudre le problème sur un horizon rallongé assez grand. On choisit d'illustrer les résultats pour un horizon rallongé de 6 ans (soit  $T = 312$  périodes d'une semaine). Bien évidemment, plus l'horizon rallongé est long, plus la politique sur la première année est fiable.

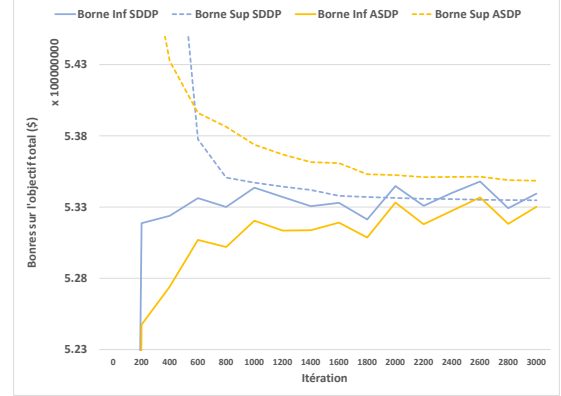
#### 6.3.1 Évolution des bornes

On exécute SDDP avec un nombre d'itérations  $I_{max} = 3000$ . À chaque 200 itérations, on calcule la borne supérieure et une estimation de la borne inférieure sur 1000 scénarios tirés aléatoirement à chaque fois (cf. section 5.7.4). Comme ASDP n'est pas une méthode itérative (une seule passe backward permet de construire une approximation des fonctions de bénéfice futur) et pour pouvoir la comparer avec SDDP, on l'exécute à chaque 200 itérations de SDDP avec un nombre d'échantillons  $D_{max}$  qui augmente de 200 à chaque fois et on calcule les bornes de la même manière et en utilisant les mêmes scénarios que SDDP.

La figure 6.7 montre l'évolution des bornes inférieure et supérieure pour les deux méthodes SDDP et ASDP.



(a)



(b)

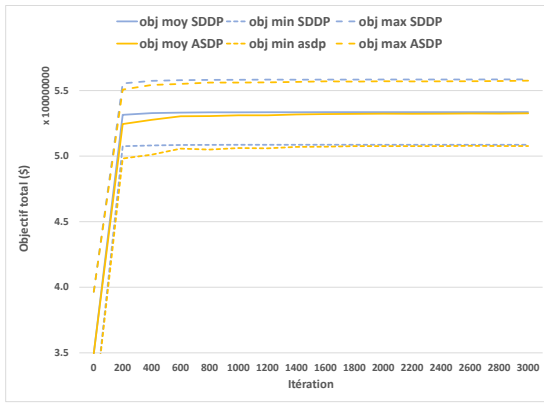
Figure 6.7 (a) évolution des bornes inférieure et supérieure ; (b) zoom sur la convergence.

Plusieurs remarques peuvent être soulignées :

- l'estimation de la borne inférieure peut fluctuer vu que les scénarios utilisés sont différents à chaque fois. À l'inverse, la borne supérieure décroît progressivement ;
- aux 400 premières itérations (ce qui équivaut à deux évaluations des bornes), la borne supérieure de ASDP décroît plus rapidement que SDDP, ce qui est prévisible vu le caractère itératif de SDDP. En effet, aux premières passes *vers l'avant* de SDDP et plus on s'éloigne dans l'horizon de planification, les coupes sont générées à des volumes bas, ce qui fait que les premières coupes véhiculent peu d'information sur le futur. On dit que SDDP nécessite une phase de *rodage* avant de commencer à échantillonner au bon endroit ;
- quant à l'estimation de la borne inférieure, on observe qu'elle est nettement meilleure dans le cas de SDDP, ceci sera justifié dans la section 6.3.3 ;
- malgré qu'elle nécessite une phase de *rodage*, SDDP converge avec un gap (entre les deux bornes) plus fin que ASDP, observons sur la figure 6.7(b) que, dans le cas de SDDP, l'estimateur de la borne inférieure est supérieur ou égal à la borne supérieure à trois reprises (les itérations 2000, 2600 et 3000).

### 6.3.2 Raffinement de la politique

Une autre métrique qui permet de comparer la performance des deux méthodes en matière de politique de gestion consiste à fixer, au tout début, un certain nombre de scénarios, par exemple 1000, et les simuler à chaque 200 itérations comme ce qu'on a fait pour les bornes. À l'issu de chaque simulation, on calcule la valeur moyenne et l'écart-type de l'objectif sur ces scénarios fixes. On rapporte sur la figure 6.8 l'évolution de la politique de gestion en matière de bénéfice sur 1000 scénarios fixes. Pour chaque méthode, SDDP et ASDP, on trace la courbe de l'objectif moyen, et ses bornes statistiques ( $\pm$  son écart-type).



(a)

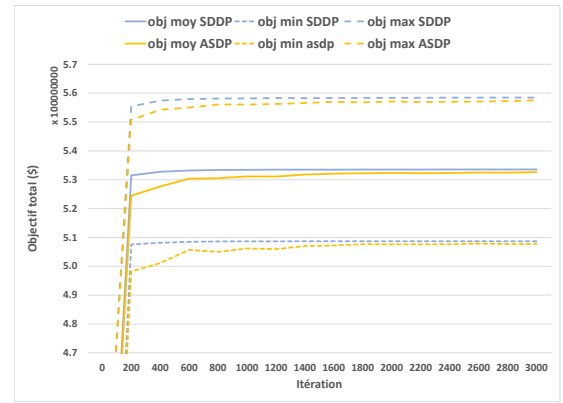
(b) *zoom*

Figure 6.8 Amélioration de la politique sur 1000 scénarios fixes

On observe que l'objectif moyen ainsi que ses bornes statistiques dans le cas de SDDP sont plus grandes que ceux dans le cas de ASDP. Ce qui confirme que la politique de gestion de SDDP est meilleure que celle de ASDP.

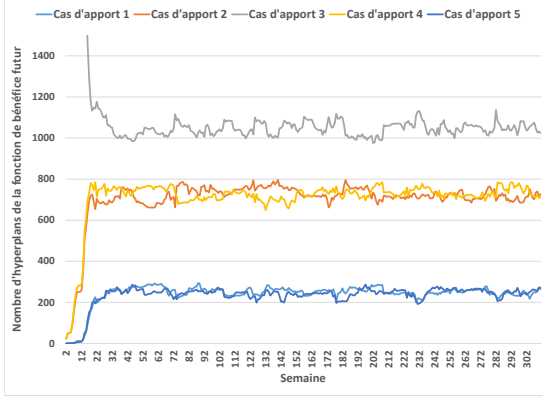
Quelles sont alors les raisons de l'avance de SDDP par rapport à ASDP ?

On essaie de répondre à cette question dans la sous-section suivante en creusant dans la manière d'échantillonner l'espace des états de l'une ou l'autre des deux méthodes.

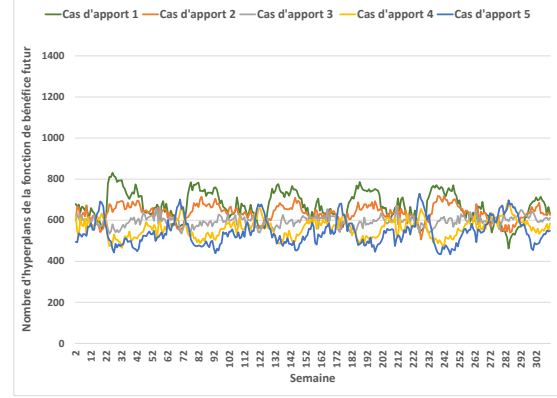
### 6.3.3 Échantillonnage de l'espace des états

L'échantillonnage de l'espace des états constitue une différence majeure entre SDDP et ASDP. La figure 6.9 montre la distribution des coupes de Benders sur les cinq cas d'apport pour

chaque semaine lorsqu'on exécute SDDP pour un nombre d'itérations  $I_{max} = 3000$  et ASDP pour un nombre d'états échantillonnés  $D_{max} = 3000$ .



(a) SDDP



(b) ASDP

Figure 6.9 Distribution des coupes de Benders sur les différents cas d'apport

On note que ASDP (b) distribue les coupes de Benders quasiment uniformément (une relativement petite fluctuation autour de la moyenne  $600 = \frac{3000}{5}$ ) sur les cinq fonctions de bénéfice futur  $\tilde{\mathcal{F}}_t^{(i)}$   $i = 1 \dots 5$  pour toutes les semaines  $t$ . Alors que SDDP (a) ne les distribue pas de cette façon. Examinons de près le processus d'apports PAR(1) discrétisé que l'on utilise illustré dans la figure 6.10



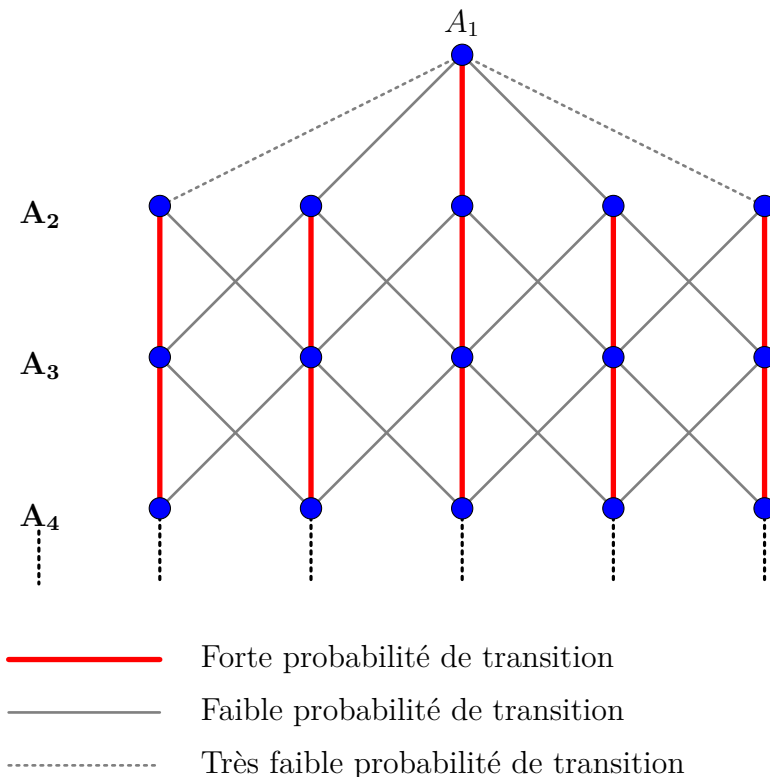


Figure 6.10 Transitions dans le processus d'apports PAR(1) discrétisé

De prime abord, on constate que le cas d'apport 3 est le plus récurrent pour toutes les semaines, suivi des deux cas d'apport 2 et 4, et en dernier les deux cas d'apport restants 1 et 5.

Vu que l'échantillonnage dans SDDP se fait par génération de scénarios aléatoires du processus d'apport discrétisé, il est donc prévu que la distribution des coupes de Benders sur les différents cas d'apport d'une semaine donnée soit plus favorable au cas d'apport 3, puis aux deux cas 2 et 4 puis aux autres cas 1 et 5. Quant à ASDP, le critère de *volume d'erreur* fait qu'elle distribue uniformément les coupes sur les cinq cas d'apport.

Ceci étant pour l'échantillonnage de l'espace de l'état *apport passé*, qu'en est-il de l'échantillonnage de l'espace de l'état *volume*?

La figure 6.11 rapporte les volumes échantillonnés de chacun des quatre réservoirs à capacité positive de LGR des deux premières années (soit 104 semaines) de l'horizon de planification rallongé. Les volumes échantillonnés par SDDP sont représentés par des points bleus, ceux échantillonnés par ASDP sont de couleur orange.

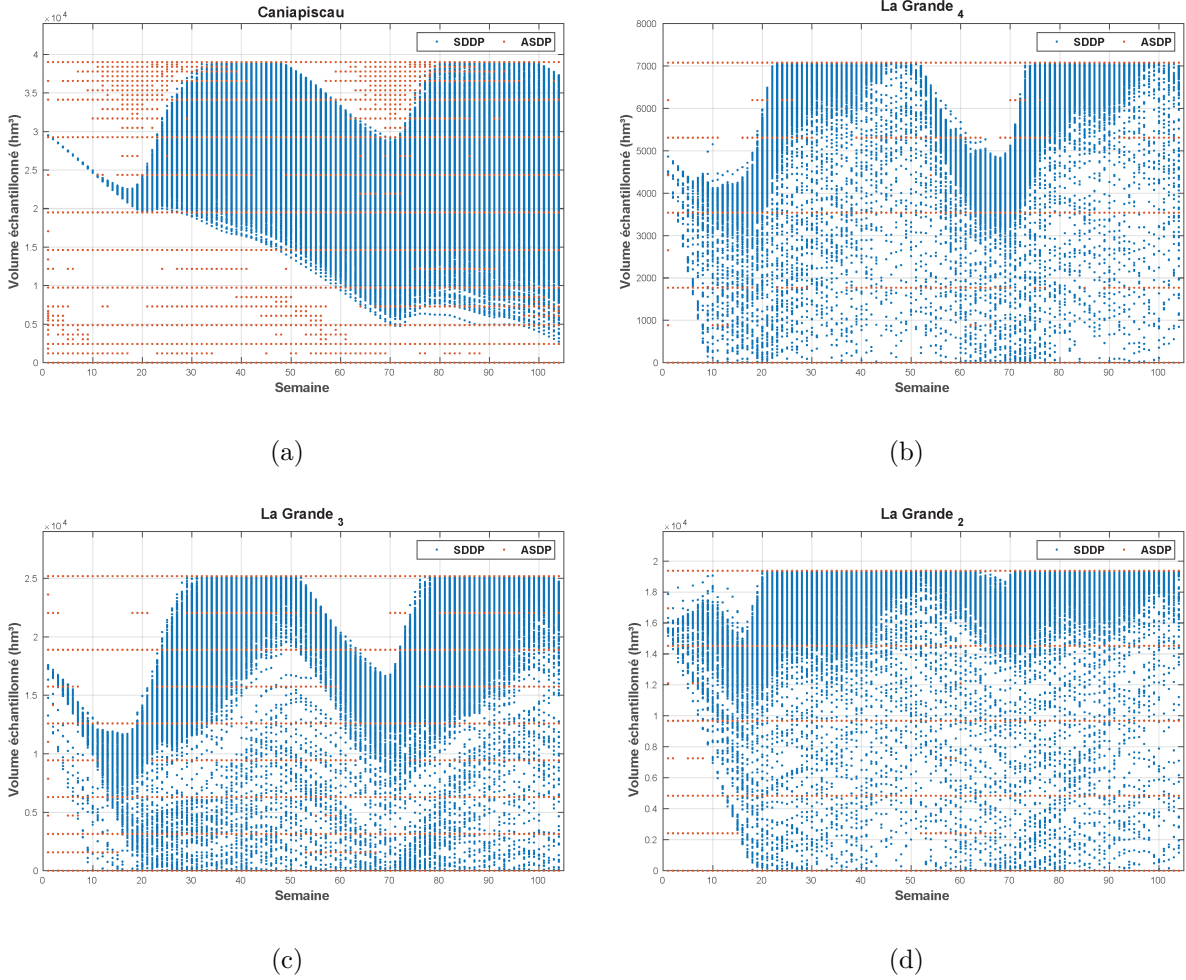


Figure 6.11 Volumes échantillonnés par les méthodes SDDP et ASDP

On voit clairement que SDDP échantillonne suivant le processus stochastique d'apport : puiser des réservoirs avant la crue de printemps pour les préparer à recevoir des quantités importantes d'apport d'eau. De plus, l'échantillonnage de SDDP contient des zones de concentration qui correspondent aux intervalles de volumes probables des semaines. On peut aussi remarquer que SDDP discrimine les réservoirs selon les performances de leurs centrales respectives : le volume du réservoir La Grande<sub>2</sub> est dans la plupart du temps échantillonné proche de sa capacité maximale vu que sa centrale est la plus performante en production (hauteur de chute de plus de 145m et une puissance maximum dépassant les 7000MW). Remarquons aussi que SDDP tient compte des volumes initiaux des réservoirs dans son échantillonnage aux premières semaines de l'horizon.

Quant à ASDP, vu son critère de découpage basé sur le *volume d'erreur* le volume qu'elle échantillonne tend à être uniforme sur toutes les semaines. On remarque aussi que ASDP peut

échantillonner à des endroits par lesquels la gestion ne va probablement pas passer (si l'on se fie au processus stochastique d'apport utilisé). Prenons, par exemple, le cas du réservoir La Grande<sub>2</sub> sur la figure 6.11(d), ASDP échantillonne son volume à des valeurs basses malgré que la performance de sa centrale augmente avec son niveau. Contrairement à SDDP, ASDP n'exploite pas les volumes initiaux des réservoirs pour restreindre son échantillonnage aux premières semaines de l'horizon.

En conclusion, SDDP échantillonne l'espace des états par apprentissage du processus stochastique d'apport et tient compte des conditions initiales du système (les volumes initiaux des réservoirs  $V_0$  et l'apport  $A_1$ ). Alors que ASDP essaie de discrétiser l'espace des états sans tenir compte de la distribution des apports et sans utiliser l'information de l'état initial du système.

#### 6.3.4 Gestion des réservoirs et de la production de LGR

On présente dans cette section la simulation de la gestion des réservoirs et de la production de LGR suite à l'application de SDDP avec un nombre d'itérations  $I_{max} = 3000$  et ASDP avec un nombre d'échantillons de l'espace des états  $D_{max} = 3000$ . La simulation est effectuée sur 1000 scénarios du processus stochastique d'apport discrétisé. La figure 6.12 donne le diagramme en boîte de ces scénarios.

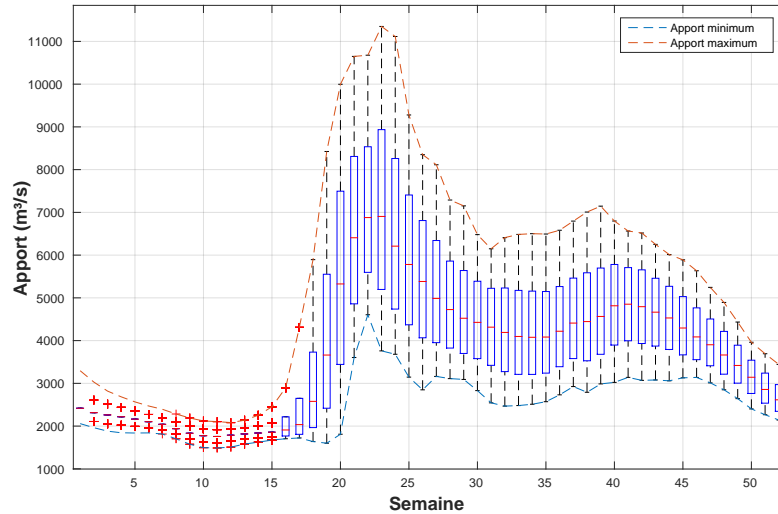
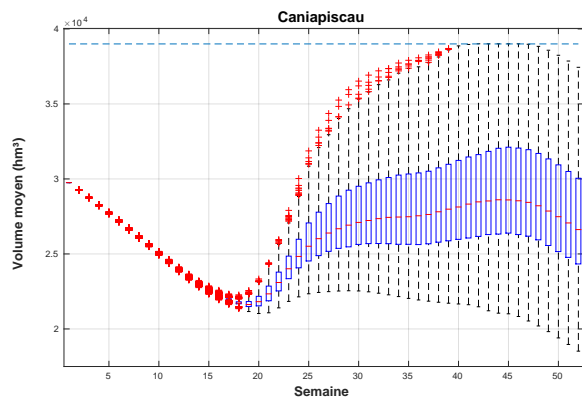


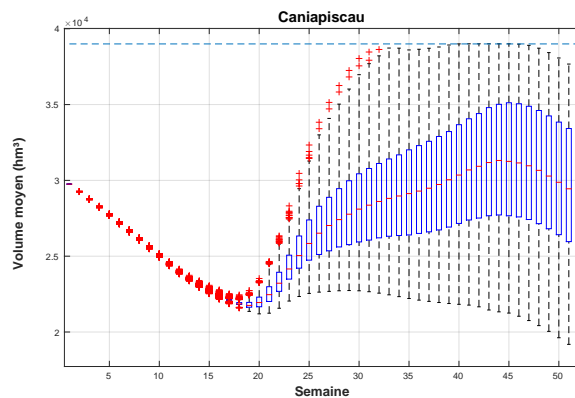
Figure 6.12 Scénarios de simulation post-politique

## Simulation de la gestion des réservoirs

Les figures 6.13, 6.14, 6.15 et 6.16 donnent les diagrammes en boîte du volume moyen simulé des réservoirs Caniapiscou, La Grande<sub>4</sub>, La Grande<sub>3</sub> et La Grande<sub>2</sub> respectivement. La figure 6.17 donne les diagrammes en boîte du débit déversé hors-système de Caniapiscou. À gauche la simulation est faite avec la politique de gestion obtenue via SDDP, à droite elle est faite avec la politique obtenue via ASDP.

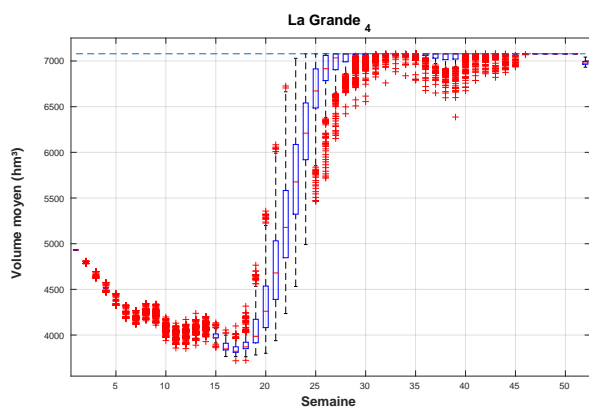


(a) SDDP

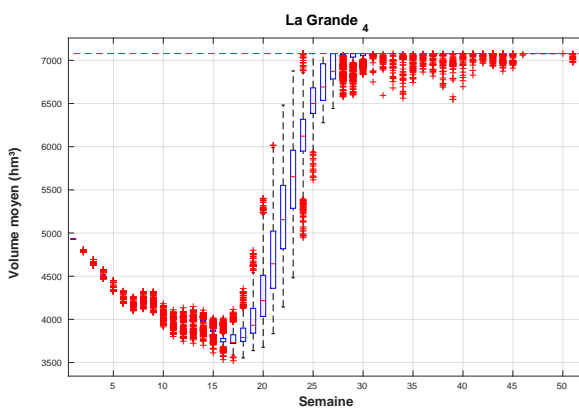


(b) ASDP

Figure 6.13 Diagramme en boîte du volume moyen simulé du réservoir Caniapiscou

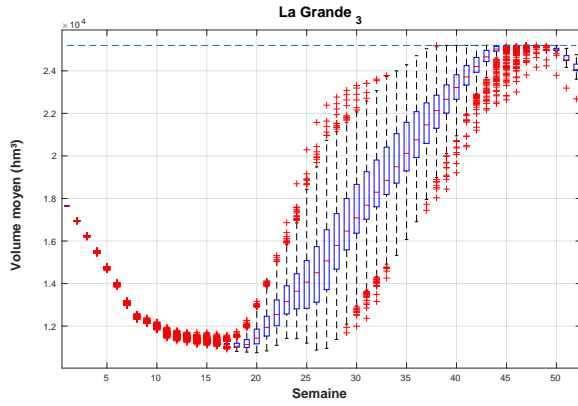


(a) SDDP

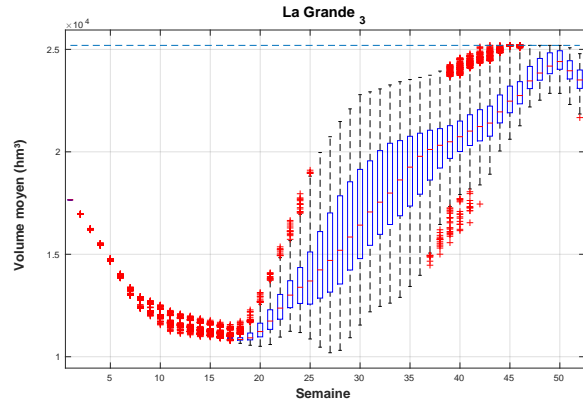


(b) ASDP

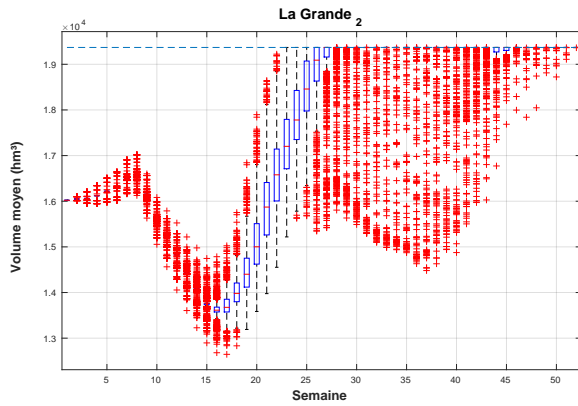
Figure 6.14 Diagramme en boîte du volume moyen simulé du réservoir La Grande<sub>4</sub>



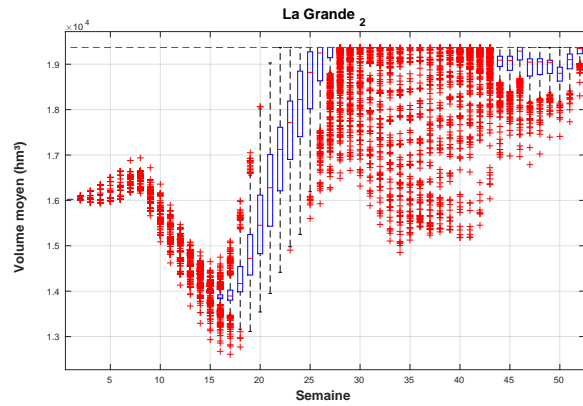
(a) SDDP



(b) ASDP

Figure 6.15 Diagramme en boîte du volume moyen simulé du réservoir La Grande<sub>3</sub>

(a) SDDP



(b) ASDP

Figure 6.16 Diagramme en boîte du volume moyen simulé du réservoir La Grande<sub>2</sub>

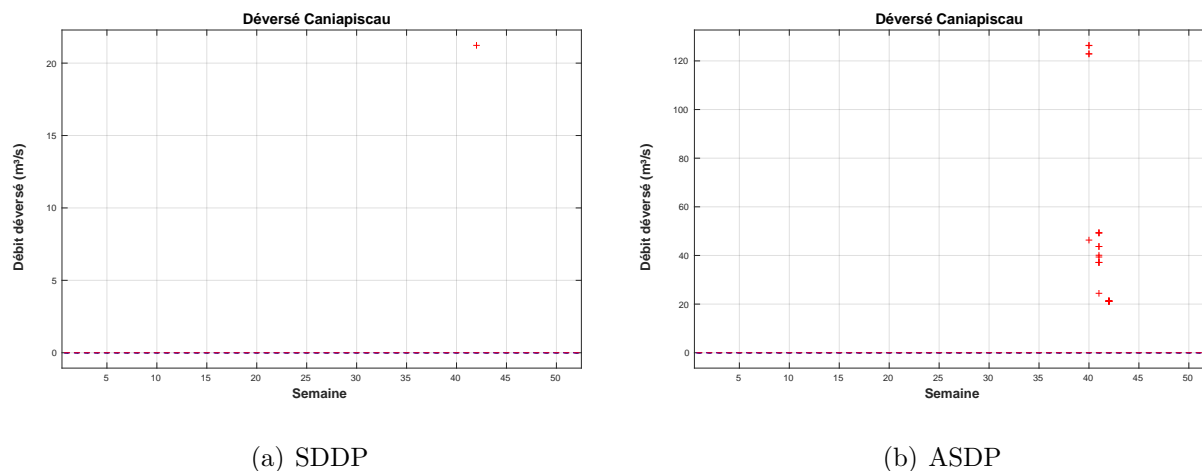


Figure 6.17 Diagramme en boîte du débit déversé simulé de l'ouvrage d'évacuation Caniapiscou

On note les différences suivantes en matière de stratégie de gestion des deux méthodes :

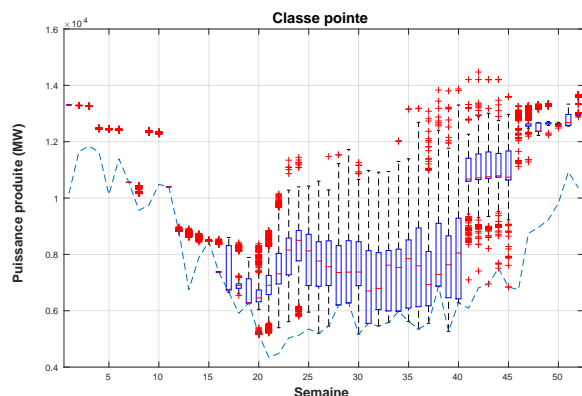
- pour le réservoir Caniapiscou, ASDP essaie de le remplir juste avant la venue de l'hiver. SDDP y garde un niveau relativement constant durant tout l'automne. Toutefois ASDP se trouve obligée de déverser hors-système plus fréquemment et en plus grandes quantités que SDDP ;
- pour le réservoir La Grande<sub>3</sub>, SDDP le remplit plus que ASDP durant la crue de printemps et jusqu'à la venue de l'hiver. Ce réservoir dispose de la 3<sup>e</sup> plus importante centrale, ainsi la stratégie de SDDP de le remplir pour se préparer à la forte demande de l'hiver est pertinente ;
- pour le réservoir La Grande<sub>2</sub>, on note une nette différence de gestion durant le début de l'hiver, SDDP le garde presque toujours à son niveau maximum ce qui est plus optimal vu que sa centrale est la plus performante.

Le comportement des deux méthodes vis-à-vis du réservoir La Grande<sub>4</sub> est relativement similaire.

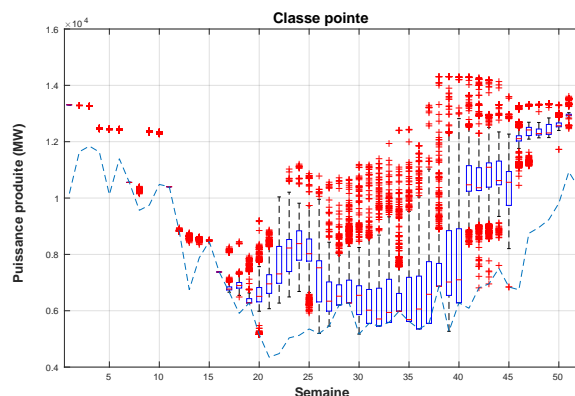
Pour ce qui est du déversé de Caniapiscou, rappelons qu'il s'agit d'un déversé hors-système du haut de LGR (cf. figure 6.1). Ce déversé est un gaspillage car s'il est plutôt turbiné par la centrale Caniapiscou, il peut être recyclé par les centrales en aval. On note que SDDP ne déverse que rarement à ce niveau. ASDP en déverse moins rarement (cf. figure 6.17).

## Simulation de la production

Les figures 6.18 et 6.19 donnent les diagrammes en boîte de la production produite simulée des classes pointe et moyenne respectivement. À gauche la simulation est faite avec la politique de gestion obtenue via SDDP, à droite elle est faite avec la politique obtenue via ASDP.

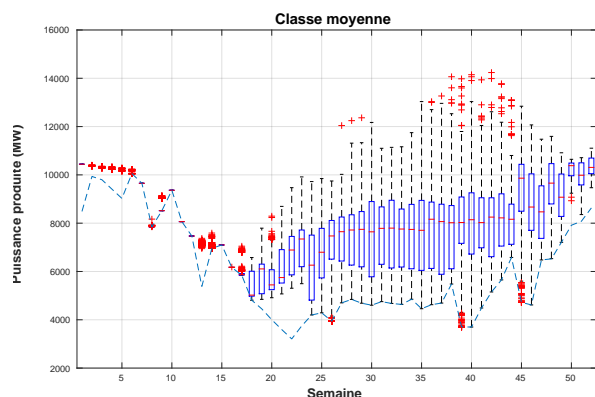


(a) SDDP

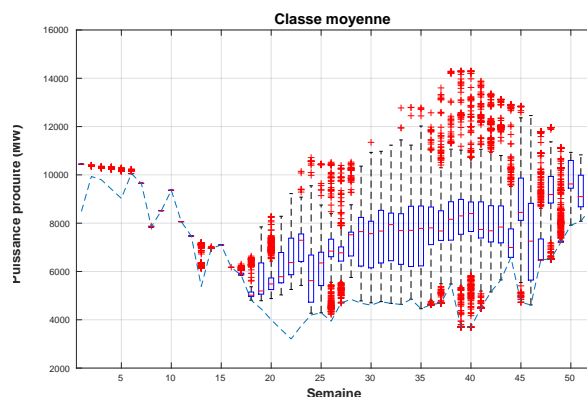


(b) ASDP

Figure 6.18 Diagramme en boîte de la puissance produite simulée durant la classe pointe



(a) SDDP



(b) ASDP

Figure 6.19 Diagramme en boîte de la puissance produite simulée durant la classe moyenne

On note que la puissance produite dans le cas de SDDP est, dans la majorité des scénarios, plus en deçà de la puissance minimum (demande) que ASDP. Ce phénomène est plus prononcé dans le cas de la classe moyenne qui dure le plus sur une semaine (75%). Ainsi la gestion

avec SDDP permet d'avoir plus de marge sur la production tout en gardant une gestion des stocks d'eau raisonnable.

#### 6.4 Retour sur les fonctions de production

On exécute SDDP avec un nombre d'itération  $I_{max} = 3000$  pour chacune des deux modélisations des fonctions de production et on rapporte son temps d'exécution. Pour la modélisation par fonctions concaves linéaires par morceaux des fonctions de production, on utilise les nombres d'hyperplans listés dans le tableau 6.8.

Tableau 6.8 Nombre d'hyperplans de l'approximation de la fonction de production

Centrale	Nombre d'hyperplans
Caniapiscau	2
La Grande <sub>4</sub>	4
La Grande <sub>3</sub>	3
La Grande <sub>2</sub>	7

On avait émis la conjecture que la modélisation par fonctions concaves linéaires par morceaux des fonctions de production (avec ses deux variantes **PartMC** et **PartMINMAX**) serait meilleure que la modélisation par enveloppe concave de deux fonctions extrêmes (**Env-Conc**) en matière de temps d'exécution de SDDP et ASDP si le nombre d'hyperplans de l'approximation est raisonnablement petit. On vérifie que c'est bien le cas puisque SDDP avec la première se termine après **25min** de temps d'exécution alors qu'avec la deuxième elle se termine après **38min**, soit un gain approximatif de **34%**.



## CHAPITRE 7 CONCLUSION

### 7.1 Synthèse des travaux

Le travail effectué et présenté dans ce mémoire peut se résumer aux points suivants

1. Modélisation d'un cas d'étude réel (La Grande Rivière) ;
2. Comparaison de différentes méthodes d'approximation concave des fonctions de production des centrales hydroélectriques et en matière de leur erreur-type (RMSE) et en matière de leur influence sur le temps d'exécution des modèles de planification étudiés ;
3. Application de deux méthodes basées sur la programmation dynamique stochastique (SDDP et ASDP) au problème de planification moyen-terme de la production d'hydro-électricité ;
4. Comparaison des deux méthodes (SDDP et ASDP) sur un cas réel et d'envergure (LGR).

Pour ce qui est des fonctions de production des centrales hydroélectriques, les deux approches proposées se sont avérées quasiment équivalentes au niveau de l'erreur-type (RMSE). La modélisation par une fonction concave linéaire par morceaux est toutefois meilleure sur le temps d'exécution des programmes linéaires utilisés dans les méthodes de planification. En plus du critère des moindres carrés utilisé par l'algorithme de partitionnement de Magnani et Boyd (2009), nous avons proposé une variante du même algorithme avec le critère de minimisation du maximum de l'erreur. Cette nouvelle variante est plus couteuse en temps d'exécution de l'algorithme de partitionnement mais nous a permis de réduire le nombre d'hyperplans tout en gardant une erreur-type (RMSE) compétitive.

L'application des méthodes SDDP et ASDP s'est faite sur LGR avec un nombre important de réservoirs. La comparaison de ses deux méthodes a montré que :

- SDDP et ASDP peuvent résoudre un problème de planification multi-réservoir ;
- SDDP échantillonne l'espace des états par simulations de scénarios tirés aléatoirement (*passé vers l'avant*). De ce fait, les coupes sont générées là où le processus stochastique des apports l'indique. ASDP n'apprend pas du processus stochastique des apports car elle n'échantillonne pas par simulations. De manière graphique, SDDP échantillonne à l'intérieur des hyper-rectangles, ASDP échantillonne les sommets des hyper-rectangles ;
- malgré qu'elle nécessite une phase de rodage de sa politique, SDDP converge vers une politique optimale après un nombre fini d'itérations ;
- ASDP distribue quasiment uniformément les coupes de Benders sur les différents cas d'ap-

port d'une semaine donnée. Cependant, ces différents cas d'apport ne sont en général pas équiprobables ;

- la convergence de SDDP est plus fine que celle de ASDP ;
- la politique obtenue via SDDP est plus proche de la politique optimale que celle obtenue via ASDP ;

## 7.2 Limitations de la solution proposée

Malgré sa convergence prouvée, SDDP nécessite une phase de *rodage* de la politique durant les premières itérations. La durée de cette phase de *rodage* dépend de la longueur de l'horizon de planification. Il s'agit de la seule limitation de SDDP.

Quant à ASDP, elle échantillonne l'espace des états parmi les sommets d'hyper-rectangles créés en temps réel. Or, la gestion des réservoirs passe par des points à l'intérieur des hyper-rectangles. Il faut donc raffiner le découpage de l'espace des états et donc couper plus encore les hyper-rectangles. On risque alors d'augmenter énormément la complexité de la méthode et la rendre proche de SDP. ASDP ne tient pas compte du processus stochastique des apports pour échantillonner l'espace des états, son caractère récursif ne permet pas de faire des simulations pour échantillonner.

## 7.3 Améliorations futures

Pour éviter les limitations de SDDP et de ASDP, nous proposons une méthode qui combine les deux ASDP-SDDP. Notons  $H$  le nombre d'années de l'horizon (rallongé) de planification, une façon de combiner les deux méthodes serait d'initialiser les coupes des  $h$  ( $h \leq H$ ) dernières années de l'horizon de planification par ASDP avec un nombre maximum d'échantillons  $D_{max}$  raisonnable (si 4 réservoirs comme le cas d'étude du chapitre 6, prendre par exemple  $D_{max} = 1000$  ou  $2000$ ), ensuite lancer SDDP sur tout l'horizon. Le cas de  $h = H$  est particulièrement intéressant. En se faisant, SDDP convergerait plus rapidement et ne nécessiterait pas de phase de *rodage*. Ainsi, on éviterait le *rodage* de SDDP et on profiterait de son échantillonnage efficace de l'espace des états.

Une amélioration possible de ASDP serait de pondérer le *volume d'erreur* par la probabilité du cas d'apport associé. Ceci permettrait un échantillonnage plus cohérent avec le processus stochastique d'apport.

## RÉFÉRENCES

- T. Aouam et Z. Yu, “Multistage stochastic hydrothermal scheduling”, dans *IEEE International Conference on Electro/Information Technology*. IEEE, 2008, pp. 66–71.
- T. Archibald, K. McKinnon, et L. Thomas, “An aggregate stochastic dynamic programming model of multireservoir systems”, *Water Resources Research*, vol. 33, no. 2, pp. 333–340, 1997.
- T. Archibald, C. Buchanan, K. McKinnon, et L. Thomas, “Nested benders decomposition and dynamic programming for reservoir optimisation”, *Journal of the Operational Research Society*, vol. 50, no. 5, pp. 468–479, 1999.
- R. E. Bellman, *Dynamic programming*. Princeton University Press, 1957.
- J. F. Benders, “Partitioning procedures for solving mixed-variables programming problems”, *Numerische Mathematik*, vol. 4, no. 1, pp. 238–252, 1962.
- D. P. Bertsekas, *Dynamic Programming and Optimal Control, Volume I*, 3e éd. Athena Scientific, 2005.
- , *Dynamic Programming and Optimal Control, Volume II*, 4e éd. Athena Scientific, 2012.
- J. R. Birge et F. Louveaux, *Introduction to stochastic programming*. Springer Science & Business Media, 2011.
- M. Breton, S. Hachem, et A. Hammadia, “A decomposition approach for the solution of the unit loading problem in hydroplants”, *Automatica*, vol. 38, no. 3, pp. 477–485, 2002.
- , “Accounting for losses in the optimization of production of hydroplants”, *IEEE Transactions on Energy Conversion*, vol. 19, no. 2, pp. 346–351, 2004.
- P.-L. Carpentier, M. Gendreau, et F. Bastin, “Long-term management of a hydroelectric multireservoir system under uncertainty using the progressive hedging algorithm”, *Water Resources Research*, vol. 49, no. 5, pp. 2812–2827, 2013.
- , “Managing hydroelectric reservoirs over an extended horizon using benders decomposition with a memory loss assumption”, *IEEE Transactions on Power Systems*, vol. 30, no. 2, pp. 563–572, 2015.

- C. Cervellera, V. C. Chen, et A. Wen, “Optimization of a large-scale water reservoir network by stochastic dynamic programming with efficient state space discretization”, *European Journal of Operational Research*, vol. 171, no. 3, pp. 1139–1151, 2006.
- C. Cervellera, A. Wen, et V. C. Chen, “Neural network and regression spline value function approximations for stochastic dynamic programming”, *Computers & Operations Research*, vol. 34, no. 1, pp. 70–90, 2007.
- V. L. de Matos, A. B. Philpott, et E. C. Finardi, “Improving the performance of stochastic dual dynamic programming”, *Journal of Computational and Applied Mathematics*, vol. 290, pp. 196 – 208, 2015.
- A. L. Diniz, F. Costa, A. L. Pimentel, L. Xavier, et M. Maceira, “Improvement in the hydro plants production function for the mid-term operation planning model in hydrothermal systems”, dans *International Conference on Engineering Optimization*. Citeseer, 2008.
- A. L. Diniz et M. E. P. Maceira, “A four-dimensional model of hydro generation for the short-term hydrothermal dispatch problem considering head and spillage effects”, *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1298–1308, 2008.
- J. Dupačová, G. Consigli, et S. W. Wallace, “Scenarios for multistage stochastic programs”, *Annals of Operations Research*, vol. 100, no. 1-4, pp. 25–53, 2000.
- M. El-Hawary et K. Ravindranath, “Hydro-thermal power flow scheduling accounting for head variations”, *IEEE Transactions on Power Systems*, vol. 7, no. 3, pp. 1232–1238, 1992.
- A. Gjelsvik, B. Mo, et A. Haugstad, “Long-and medium-term operations planning and stochastic modelling in hydro-dominated power systems based on stochastic dual dynamic programming”, dans *Handbook of Power Systems I*. Springer, 2010, pp. 33–55.
- Q. Goor, C. Halleux, Y. Mohamed, et A. Tilmant, “Optimal operation of a multipurpose multireservoir system in the eastern Nile river basin”, *Hydrology and Earth System Sciences*, vol. 14, no. 10, pp. 1895–1908, 2010.
- Q. Goor, R. Kelman, et A. Tilmant, “Optimal multipurpose-multireservoir operation model with variable productivity of hydropower plants”, *Journal of Water Resources Planning and Management*, vol. 137, no. 3, pp. 258–267, 2010.
- T. Halliburton *et al.*, “An optimal hydrothermal planning model for the New Zealand power system”, *Australian Journal of Electrical & Electronics Engineering*, vol. 1, no. 3, p. 193, 2004.

A. Hammadia, “Contributions à l’optimisation, en temps réel et à court terme, des ressources hydroélectriques d’une rivière”, Thèse de doctorat, École Polytechnique de Montréal, Montréal, QC, Canada, 2000.

T. Homem-de Mello, V. L. de Matos, et E. C. Finardi, “Sampling strategies and stopping criteria for stochastic dual dynamic programming : a case study in long-term hydrothermal scheduling”, *Energy Systems*, vol. 2, no. 1, pp. 1–31, 2011.

W.-C. Huang, R. Harboe, et J. J. Bogardi, “Testing stochastic dynamic programming models conditioned on observed or forecasted inflows”, *Journal of Water Resources Planning and Management*, vol. 117, no. 1, pp. 28–36, 1991.

S. A. Johnson, J. R. Stedinger, C. A. Shoemaker, Y. Li, et J. A. Tejada-Guibert, “Numerical solution of continuous-state dynamic programs using linear and spline interpolation”, *Operations Research*, vol. 41, no. 3, pp. 484–500, 1993.

C. R. P. Jr et P. K. Kitanidis, “Limitations of deterministic optimization applied to reservoir operations”, *Journal of Water Resources Planning and Management*, vol. 125, no. 3, pp. 135–142, 1999.

M. Karamouz et M. H. Houck, “Comparison of stochastic and deterministic dynamic programming for reservoir operating rule generation”, *Journal of the American Water Resources Association*, vol. 23, no. 1, pp. 1–9, 1987.

M. Karamouz et H. V. Vasiliadis, “Bayesian stochastic optimization of reservoir operation using uncertain forecasts”, *Water Resources Research*, vol. 28, no. 5, pp. 1221–1232, 1992.

M. Karamouz, M. H. Houck, et J. W. Delleur, “Optimization and simulation of multiple reservoir systems”, *Journal of Water Resources Planning and Management*, vol. 118, no. 1, pp. 71–81, 1992.

S. Krau, G. Émiel, J. Merleau, et R. Leconte, “Adaptive discretization of the state space for stochastic dynamic programming applied to multi-reservoir system”, 2015, Manuscrit non publié.

T. Kristiansen, “Financial risk management in the electric power industry using stochastic optimization”, *Advanced Modeling and Optimization*, vol. 6, no. 2, pp. 17–24, 2004.

J. W. Labadie, “Optimal operation of multireservoir systems : state-of-the-art review”, *Journal of Water Resources Planning and Management*, vol. 130, no. 2, pp. 93–111, 2004.

- M. E. P. Maceira, V. S. Duarte, D. D. J. Penna, L. A. M. Moraes, et A. C. G. Melo, “Ten years of application of stochastic dual dynamic programming in official and agent studies in brazil-description of the newave program”, *16th PSCC, Glasgow, Scotland*, pp. 14–18, 2008.
- A. Magnani et S. P. Boyd, “Convex piecewise-linear fitting”, *Optimization and Engineering*, vol. 10, no. 1, pp. 1–17, 2009.
- B. Mo, A. Gjelsvik, et A. Grundt, “Integrated risk management of hydro power scheduling and contract management”, *IEEE Transactions on Power Systems*, vol. 16, no. 2, pp. 216–221, 2001.
- M. V. F. Pereira et L. M. V. G. Pinto, “Application of decomposition techniques to the mid-and short-term scheduling of hydrothermal systems”, *IEEE Transactions on Power Apparatus and Systems*, no. 11, pp. 3611–3618, 1983.
- M. Pereira, N. Campodonico, et R. Kelman, “Application of stochastic dual dp and extensions to hydrothermal scheduling”, *PSRI Technical Report*, 1999.
- M. V. F. Pereira et L. M. V. G. Pinto, “Multi-stage stochastic optimization applied to energy planning”, *Mathematical Programming*, vol. 52, no. 1-3, pp. 359–375, 1991.
- A. B. Philpott et V. L. de Matos, “Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion”, *European Journal of Operational Research*, vol. 218, no. 2, pp. 470–483, 2012.
- A. B. Philpott et Z. Guan, “On the convergence of stochastic dual dynamic programming and related methods”, *Operations Research Letters*, vol. 36, no. 4, pp. 450–455, 2008.
- R. J. Pinto, C. L. Borges, et M. E. Maceira, “An efficient parallel algorithm for large scale hydrothermal system operation planning”, *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4888–4896, 2013.
- D. Rani et M. M. Moreira, “Simulation–optimization modeling : a survey and potential application in reservoir systems operation”, *Water Resources Management*, vol. 24, no. 6, pp. 1107–1138, 2010.
- T. Røtting, A. Gjelsvik *et al.*, “Stochastic dual dynamic programming for seasonal scheduling in the norwegian power system”, *IEEE Transactions on Power Systems*, vol. 7, no. 1, pp. 273–279, 1992.

- M. Saad, A. Turgeon, P. Bigras, et R. Duquette, “Learning disaggregation technique for the operation of long-term hydroelectric power systems”, *Water Resources Research*, vol. 30, no. 11, pp. 3195–3202, 1994.
- T. N. D. Santos et A. L. Diniz, “A new multiperiod stage definition for the multistage benders decomposition approach applied to hydrothermal scheduling”, *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1383–1392, 2009.
- A. Shapiro, “Analysis of stochastic dual dynamic programming method”, *European Journal of Operational Research*, vol. 209, no. 1, pp. 63–72, 2011.
- J. R. Stedinger, B. F. Sule, et D. P. Loucks, “Stochastic dynamic programming models for reservoir operation optimization”, *Water Resources Research*, vol. 20, no. 11, pp. 1499–1505, 1984.
- J. A. Tejada-Guibert, S. A. Johnson, et J. R. Stedinger, “Comparison of two approaches for implementing multireservoir operating policies derived using stochastic dynamic programming”, *Water Resources Research*, vol. 29, no. 12, pp. 3969–3980, 1993.
- , “The value of hydrologic information in stochastic dynamic programming models of a multireservoir system”, *Water Resources Research*, vol. 31, no. 10, pp. 2571–2579, 1995.
- A. Tilmant, L. Beevers, et B. Muyunda, “Restoring a flow regime through the coordinated operation of a multireservoir system : The case of the zambezi river basin”, *Water Resources Research*, vol. 46, no. 7, 2010.
- A. Tilmant et R. Kelman, “A stochastic approach to analyze trade-offs and risks associated with large-scale water resources systems”, *Water resources research*, vol. 43, no. 6, 2007.
- A. Turgeon, “Optimal operation of multireservoir power systems with stochastic inflows”, *Water Resources Research*, vol. 16, no. 2, pp. 275–283, 1980.
- , “A decomposition method for the long-term scheduling of reservoirs in series”, *Water Resources Research*, vol. 17, no. 6, pp. 1565–1570, 1981.
- , “Stochastic optimization of multireservoir operation : The optimal reservoir trajectory approach”, *Water Resources Research*, vol. 43, no. 5, 2007.
- A. Turgeon et R. Charbonneau, “An aggregation-disaggregation approach to long-term reservoir management”, *Water Resources Research*, vol. 34, no. 12, pp. 3585–3594, 1998.

S. Yakowitz, “Dynamic programming applications in water resources”, *Water Resources Research*, vol. 18, no. 4, pp. 673–696, 1982.

W. W.-G. Yeh, “Reservoir management and operations models : A state-of-the-art review”, *Water Resources Research*, vol. 21, no. 12, pp. 1797–1818, 1985.



## ANNEXE A REPRÉSENTATION DES FONCTIONS CONCAVES LINÉAIRES PAR MORCEAUX DANS LES MODÈLES LINÉAIRES

Dans cet annexe, on expose plusieurs approches équivalentes pour représenter une fonction concave linéaire par morceaux dans un modèle linéaire de maximisation.

Considérons une fonction concave linéaire par morceaux  $x \mapsto y$  (figure A.1)

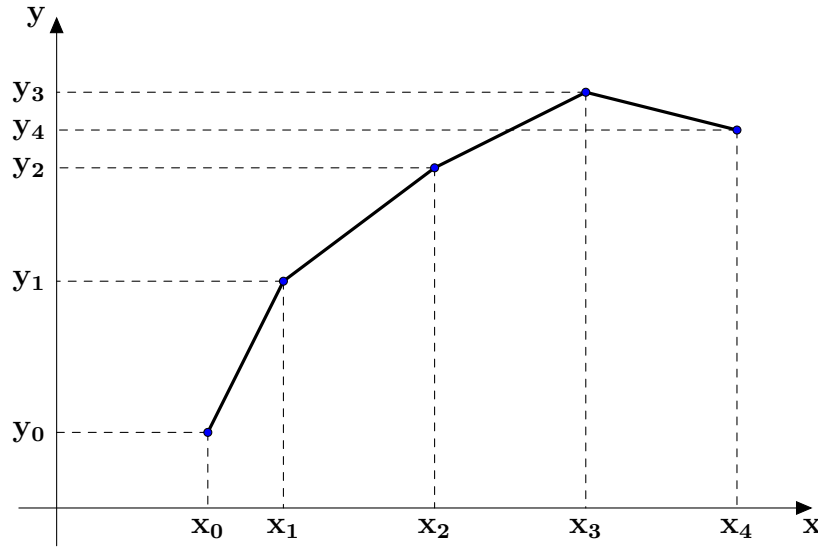


Figure A.1 Une fonction concave linéaire par morceaux

On a pour  $x_0 \leq x \leq x_4$

$$\begin{aligned}
 y(x) &= \max_{\lambda} \sum_{i=0}^4 \lambda_i y_i &= y_0 + \max_{\mathbf{sv}} \sum_{j=1}^4 a_j \mathbf{sv}_j &= \max_{\mathbf{z}} \mathbf{z} \\
 \text{s.à } \sum_{i=0}^4 \lambda_i x_i &= x &\text{s.à } \sum_{j=1}^4 \mathbf{sv}_j &= x - x_0 &\text{s.à } \mathbf{z} &\leq a_j x + b_j, \quad j = 1 \dots 4 \\
 \sum_{i=0}^4 \lambda_i &= 1 &\mathbf{sv}_j &\leq x_j - x_{j-1}, \quad j = 1 \dots 4 \\
 \lambda &\geq 0 &\mathbf{sv} &\geq 0
 \end{aligned}$$

Où  $a_j$  et  $b_j$  sont respectivement la pente et l'ordonnée à l'origine du  $j^{\text{e}}$  plan affine correspondant au  $j^{\text{e}}$  morceau.

Si, dans un modèle de programmation linéaire complexe, on utilise deux variables  $Y$  et  $X$  et

que l'on souhaite que  $Y = y(X)$ , il faut s'assurer que pour une valeur fixe de la variable  $X$  le modèle va aller chercher la plus grande valeur possible de  $Y$ . Si c'est le cas il suffit alors d'incorporer un des ensembles de contraintes (équivalents) parmi les suivants

$$\left\{ \begin{array}{l} Y = \sum_{i=0}^4 \lambda_i y_i \\ X = \sum_{i=0}^4 \lambda_i x_i \\ \sum_{i=0}^4 \lambda_i = 1 \\ \lambda \geq 0 \end{array} \right. \equiv \left\{ \begin{array}{l} Y = y_0 + \sum_{j=1}^4 a_j \mathbf{sv}_j \\ X = x_0 + \sum_{j=1}^4 \mathbf{sv}_j \\ \mathbf{sv}_j \leq x_j - x_{j-1}, \quad j = 1 \dots 4 \\ \mathbf{sv} \geq 0 \end{array} \right. \equiv \left\{ Y \leq a_j X + b_j, \quad j = 1 \dots 4 \right.$$

Dans le modèle de planification de la production d'hydroélectricité qu'on traite dans ce mémoire, il est question de maximiser la production totale. Ainsi,

- les contraintes de type  $\mathbf{p} = p_c(\mathbf{q})$  sont exactes ;
- les contraintes de type  $\mathbf{l}_r^u = l_r(\mathbf{v})$  sont aussi exactes car, à un débit donné, plus le niveau du réservoir  $r$  est grand plus la centrale associée à  $r$  produit d'électricité.

**Remarque :**

Le développement précédent s'est fait pour un nombre de morceaux égal à 4. Il reste généralisable à un nombre de morceaux quelconque.